NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

IMPROVING THE INTEGRATED TRAINING CENTER (ITC) MODEL TO ACHIEVE MORE ACCURATE TIME TO TRAIN ESTIMATES

by

Joe K. Dummar

June 2011

Thesis Advisor:	Chad Seagren
Second Reader:	Thomas W. Lucas

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2011	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE Improving the Integrated Training Center (ITC) Model to Achieve More Accurate Time to Train Estimates			**5. FUNDING NUMBERS**
6. AUTHOR(S) Joe K. Dummar			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number _N/A_.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. Distribution Code** A

13. ABSTRACT (maximum 200 words)

The F-35B Joint Strike Fighter is critical to the future of Marine Corps fixed-wing aviation. The ability to man the Joint Strike Fighter with properly trained pilots is as important as the development and acquisition of the aircraft itself. This thesis examines the Integrated Training Center (ITC) model, which simulates the training process in order to determine expected time to train for Marine pilots trained at Eglin Air Force Base, Florida. First, we demonstrate that weather is an important factor in the Integrated Training Center model. Legal and financial constraints restrict modification of the ITC model, so we create a replica that is statistically similar to the flight management process found in that model. We then modify the replica model to more accurately reflect reality. This improved model uses a continuous time Markov process to model weather. We show that the independent weather assumption in the ITC model is inappropriate. We recommend modifying the ITC model to reflect good weather as a resource necessary to conduct a flight, and one that is only intermittently available. Ultimately, these tools will be provided to decision makers so that they may implement these changes to make better decisions.

14. SUBJECT TERMS Joint Strike Fighter, Integrated Training Center, Weather, Pilot Training, Time to Train, Continuous Time Markov Process, Nearly Orthogonal Latin Hypercube			**15. NUMBER OF PAGES** 99
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

IMPROVING THE INTEGRATED TRAINING CENTER (ITC) MODEL TO ACHIEVE MORE ACCURATE TIME TO TRAIN ESTIMATES

Joe K. Dummar
Captain, United States Marine Corps
B.S., Hawaii Pacific University, 2005

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 2011

Author: Joe K. Dummar

Approved by: Chad Seagren
Thesis Advisor

Thomas W. Lucas
Second Reader

Robert F. Dell
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The F-35B Joint Strike Fighter is critical to the future of Marine Corps fixed-wing aviation. The ability to man the Joint Strike Fighter with properly trained pilots is as important as the development and acquisition of the aircraft itself. This thesis examines the Integrated Training Center (ITC) model, which simulates the training process in order to determine expected time to train for Marine pilots trained at Eglin Air Force Base, Florida. First, we demonstrate that weather is an important factor in the Integrated Training Center model. Legal and financial constraints restrict modification of the ITC model, so we create a replica that is statistically similar to the flight management process found in that model. We then modify the replica model to more accurately reflect reality. This improved model uses a continuous-time Markov process to model weather. We show that the independent weather assumption in the ITC model is inappropriate. We recommend modifying the ITC model to reflect good weather as a resource necessary to conduct a flight, and one that is only intermittently available. Ultimately, these tools will be provided to decision makers so that they may implement these changes to make better decisions.

THIS PAGE INTENTIONALLY LEFT BLANK

THESIS DISCLAIMER

The reader is cautioned that the computer programs presented in this research may not have been exercised for all cases of interest. While every effort has been made, within the time available, to ensure that the programs are free of computational and logical errors, they cannot be considered validated. Any application of these programs without additional verification is at the risk of the user.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

.csv	Comma Separated Value File Format
A/C	Aircraft
AFB	Air Force Base
AICc	Corrected Akaike Information Criterion
ANOVA	Analysis of Variance
COA	Course of Action
CQ	Carrier Qualification
DES	Discreet-Event Simulation
DOE	Design of Experiments
DWQ	Mean Delay in Wait Queue
FMS	Full Mission Simulator
GUI	Graphical User Interface
IP	Instructor Pilot
ITC	Integrated Training Center
JSF	Joint Strike Fighter
LANTIRN	Low-Altitude Navigation and Targeting Infra-Red for Night
METAR	Meteorological Terminal Aviation Routine Weather Report
MODSIM	Model Simulator Software
MOE	Measure of Effectiveness
MRT	Mission Readiness Trainer
NOLH	Nearly Orthogonal Latin Hypercube
NPS	Naval Postgraduate School
OLS	Ordinary Least Squares
TTT	Time to Train

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The F-35B Joint Strike Fighter (JSF) is a critical step in the future of the Marine Corps' fixed-wing aviation. This aircraft (A/C) will make a significant difference in combat effectiveness and its success will play a vital role in the future of the Marine Corps. The ability to man the JSF with properly trained pilots is as important as the development and acquisition of the A/C itself.

Time to Train (TTT) estimates have a major influence on manpower requirements for pilot training, as instructor pilots (IPs) and trainees are a large part of the costs involved in the training cycle. If TTT estimates are inaccurate, there could be considerable delays that negatively impact the manning of operational squadrons. These delays could adversely affect the combat effectiveness of the Marine Corps (Rabachault, 2011).

The Marine Corps needs to understand the length of time required to train pilots in order to appropriately assign resources to this process and make accurate plans. The Integrated Training Center (ITC) model simulates the training process in order to determine TTT (Kenny, 2010). The ITC model is useful, but contains many flaws that significantly impact the training estimates.

TTT estimates depend upon many factors, both controllable and uncontrollable. The ITC model takes a set of input data and simulates the induction, training, and graduation of student pilots until all pilots have completed their respective courses. Student TTT is one output of the ITC model, which assesses the training capacity and acquisition requirements for the first JSF ITC currently located at Eglin Air Force Base, Florida.

The ITC model is essentially a series of queues that approximate the resources students utilize in the training process and accounts for the time students spend actually using those resources and waiting to use them. Students are serviced by various training resources such as classrooms, A/C, and simulators. Students progress through the training process using those available resources. If resources are not available, they go

into waiting queues until the resources are on hand. The model records student entry and exit times as object attributes for both students and resources (Kenney, 2009). The ITC currently uses an independent Bernoulli trial, similar to a dice roll, to approximate flight cancellations due to weather. This could result in one student's flight being cancelled within a few minutes of another pilot taking off.

The ITC model is written in Model Simulator Software (MODSIM III) and is proprietary software; therefore, the user cannot make modifications. In order to analyze the effects of weather changes on the model, it is necessary to replicate the ITC model's use of weather. This thesis uses Simkit, a discrete-event simulation software, to replicate the weather portion of the ITC model. It also uses Simkit to simulate the weather in a different model using a stochastic process, which will hold up entire groups of students during bad weather. This thesis takes a more realistic approach when modeling weather-related flight cancellations.

This thesis investigates the ITC model through experimental design and regression analysis. Summer and winter cancellation rates are some of the most important factors in the current model. This research confirms that weather is important in the ITC model and that small changes in bad weather rates using the current model can result in significant delays in TTT.

A replica model, which is statistically similar to the flight management process found in the ITC model, is created in Simkit. This model is necessary because we cannot modify the ITC model directly. We analyze the weather piece and verify that summer and winter cancellation rates are important in this model. Then, we analyze resource utilization rates to ensure we are not overusing resources. No resource limitations are placed on this model helping in the analysis of A/C usage to determine if usage levels are acceptable or not. This creates an issue, in that the estimates shown here are lower bounds. This thesis shows that slight changes in weather can significantly impact the TTT for each pilot.

An improved model, similar to the replica model, is also created in Simkit. This model contains a continuous-time Markov process that models the state of the weather.

In this model, the weather is treated as a resource that is only intermittently available. The process whereby the weather changes approximate the four seasons is flexible enough to model the weather of most geographical locations. This model is also analyzed for resource utilization. This resource utilization is found to be much more volatile than in the replica model. This resource utilization can greatly impact the TTT estimates generated using the ITC model. Finally, with this model, we demonstrate that modeling the effect of weather on flights as independent Bernoulli trials is inappropriate and leads to inaccurate training time estimates.

We recommend modifying the ITC model to reflect good weather as a resource, necessary to conduct a flight, and one that is only intermittently available. One way to do this is to use a continuous-time Markov process, as we demonstrate in this thesis.

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

First of all, I must thank my beautiful and extremely supportive wife, Michelle. You pushed me to move forward even when it seemed, at times, impossible. Without your love, understanding, and support there is no way I could have devoted the time necessary to complete a quality product. I will be forever grateful to you for your love and encouragement. I need also thank my three wonderful boys, who helped me keep perspective through it all.

I want to thank Chad Seagren for his vital help in developing the models, and his advice all throughout the writing process. Without his guidance I would never have made it through this time.

I need to thank Paul Roeder and Steve Upton, who were instrumental in running my model, developing the designs of experiments, and finally, in analyzing my results. Their help was a significant factor in my ability to complete this analysis.

I also want to thank Professor Tom Lucas and Colonel Ed Lesnowicz, USMC (Ret.). Thank you for providing the topic, helping me with the modeling and analysis, supporting my efforts through the write-up, and keeping me on track throughout the entire process.

Finally, I would like to say thank you to Richard Mastowski; your editorial services were incredibly constructive and helped bring this thesis together.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

> If the F-35B is canceled, the service has no other option for replacing its aging Harrier jets capable of operating off smaller ships. . . . That would eventually leave the nation with half as many ships deploying fighter jets. . . . There is not a plan B. . . . This is bigger than the Marine Corps, [General Amos] concluded.
>
> —G.C. Kovach, 2010, p. 1

A. OVERVIEW

The F-35B Joint Strike Fighter (JSF) is a critical step in the future of Marine Corps' fixed-wing aviation. This aircraft (A/C) will make a significant difference in combat effectiveness and its success will play a vital role in the future of the Marine Corps. The ability to man the JSF with properly trained pilots is as important as the development and acquisition of the A/C itself.

In January 2011, Defense Secretary Robert Gates expressed concern with the increasing costs associated with the F-35 program when he said, "The culture of endless money that has taken hold must be replaced by a culture of restraint." He singled out the Marine Corps' variant F-35B and ordered "a two-year probation," saying it "should be canceled" if corrections are unsuccessful (Rabachault, 2011, p. 1).

With an estimated budget of over $380 billion, the JSF is one of the most expensive procurement programs ever in the Department of Defense. Time to Train (TTT) estimates have a major influence on manpower requirements for pilot training, as instructor pilots (IPs) and trainees are a large part of the costs involved in the training cycle and need to be addressed properly. If underestimates are not addressed early, then there will be considerable delays that negatively impact the manning of operational squadrons. These delays will also have adverse effects on the combat effectiveness of the Marine Corps (Rabachault, 2011).

The Marine Corps needs to understand the length of the training cycle in order to effectively train F-35B pilots. The Integrated Training Center (ITC) model simulates the training process in order to determine TTT (Kenny, 2010). The ITC model is useful, but

there are many flaws with it that may significantly impact the training estimates it produces for the Marine Corps (Lucas, 2010).

B. BACKGROUND

The JSF program is the Department of Defense's focal point for the next-generation strike A/C weapon system for the Navy, Air Force, Marine Corps, and our allies. The JSF has three variants: the F-35A Conventional Takeoff and Landing A/C used by the Air Force, the F-35B Short Takeoff and Vertical Landing used by the Marine Corps, and the F-35C used by the Navy for Carrier Launch and Recovery. This research focuses on the F-35B.

1. Training Process

The ITC model takes a set of input data and simulates the induction, training, and graduation of student pilots until all of them have completed their respective courses. Student TTT is an output of the ITC model. TTT depends upon the syllabus written into the ITC model or the information changed directly in the Graphical User Interface (GUI). Delays to the training process occur when resources are not available for use. The ITC model assesses the training throughput and capacity for the first JSF ITC located at Eglin Air Force Base (AFB), Florida, as well as for other training facilities abroad. It is also being used as a guide for resource acquisitions for the training center.

2. ITC Model

The ITC model is a Monte Carlo queuing-based model that uses uniform distributions as a basis for all random behavior. The model is implemented in Model Simulator Software (MODSIM) III, which is software that enables an object-oriented, discrete-event simulation framework (Goble, 1997). Queues of students are serviced by various training resources such as classrooms, A/C, and simulators. Students progress through the training process using the available resources. If resources are not available, they go into waiting queues until the resources are on hand. The queues use a first-in-first-out system. Wait queues hold students who are waiting to train, while resources are held by resource queues that are available for use. Student entry and exit

times are recorded as object attributes for both students and resources. The following resources are accounted for using the ITC model (Kenney, 2009):

- **Training Devices:** Devices such as the Full Mission Simulators (FMS) help students to gain a better understanding of the A/C before stepping into the cockpit.
- **Classrooms:** Rooms are set apart and can accommodate up to 12 pilots per class. These rooms are used mostly for annual training and some general knowledge instruction.
- **Self-Paced, Computer-Based Training Work Stations:** These are only available if a space is available. Most of the preflight instruction is based upon these programs. Once a certain level is reached, an event-based test is given to ensure the learning has taken place.
- **Mission Brief/Debrief Rooms:** These are smaller rooms to be utilized before and after training flights or FMS flights, in order to allow feedback to pilots on areas for improvement.
- **Military IPs:** IPs are the instructors who teach the pilots as they go through their training. They are limited to the number of training A/C available. For new A/C, like the F-35B, the IPs also need training.
- **Military Maintainer Instructors:** These instructors are limited to approximately 30 hours of training per week, with another 10 hours set aside for counseling, syllabus and course content review, lesson preparation, and other instructor-related duties.
- **Contract Academic IPs:** Contracted instructors who will be relied upon for expertise and to conduct a majority of the training during the start of the program until IPs are able to be created from within.
- **Training A/C:** A/C purchased specifically for use at the training facility.
- **Runways and Training Airspace:** These resources, such as military operating areas, low-level routes, scored target ranges, etc., will have availability only during specific times. They will be utilized as necessary when pilots are ready to fly the actual A/C.

The ITC currently uses an independent Bernoulli trial to simulate weather. As a student enters the flight queue a random number is generated, and if that number falls below the weather cancellation rate, he is not allowed to fly. Each student that enters the flight queue is tested independently of anyone else. This could result in student A being cancelled for weather and a few minutes later student B is allowed to fly (Lucas, 2010).

3. Discreet-Event Simulation (DES) and Simkit

Weather is not represented correctly in the ITC model. The model is written in MODSIM III and is proprietary software; therefore, the user cannot modify it. In order to analyze the effects of weather changes on the model, it is necessary to replicate the ITC model's use of weather in the flight event. Simkit is a DES modeling language that was developed at the Naval Postgraduate School. It uses the Java programming language for all its object-oriented programming (Buss, 2001). This thesis uses Simkit to replicate the weather portion of the ITC model. It also uses Simkit to simulate an improved weather model using a stochastic approach to weather. These two models create a way of comparing improvements to the weather model. This thesis goes into further details about the Simkit models in Chapters III and IV.

C. OBJECTIVES

Using design of experiments (DOE) and simulation, this thesis studies the ITC model and the impact of weather on the expected TTT for Marine pilots being trained at Eglin AFB. This thesis analyzes the effects of extending the model to more realistically account for weather. These changes will greatly improve the ability of the model to give better TTT estimates.

This thesis confirms that weather is important in the ITC model. It describes a Simkit model that is statistically similar to the flight event process found in the ITC model. It also develops a Simkit model that more accurately models weather. The assumption of weather being modeled as an independent Bernoulli trial is inappropriate and leads to inaccurate TTT estimates. We recommend modifying the ITC model to reflect weather as a resource, necessary to conduct a flight, and one that is only intermittently available. Ultimately, these tools will be provided to decision makers so that they may implement these changes to make better decisions, based upon more accurate TTT estimates.

D. RESEARCH QUESTIONS

First, how significant is weather to the current model? Second, how sensitive are TTT estimates with respect to weather? Finally, a comparison of the proposed Simkit model to the current model is done to see if the proposed changes are statistically significant. This comparison is done using linear regression analysis, which allows a quantitative comparison of distinct models.

E. SCOPE, LIMITATIONS, AND DATA ASSUMPTIONS

1. Scope

The scope of this thesis is to determine the significance of weather on the current model, in order to improve the TTT estimates that the ITC model produces. We do this using robust analysis in order to decide whether the difference between the current model estimates are statistically different from estimates using the improved model.

2. Limitations

The most important limitation of this thesis is that researchers cannot change the ITC model. The model is written in MODSIM III and the model itself is proprietary software. Changing the model requires the Marine Corps to incur excess costs. The Marine Corps can either pay a contractor to make changes, or buy the rights to the model and make the changes themselves. Funds for doing so, however, are not available at this time.

3. Model Assumptions

a. ITC Model Assumptions

This thesis assumes the latest plan for resource allocation is up-to-date and accurate. The plan is written into Course of Action (COA) 4, which consists of 14 syllabai, A/C induction schedules, IP induction schedules, and other basic resources required for this model to run. Using the latest COA gives one the opportunity to study

TTT according to the most recent resource allocation numbers provided by the Marine Corps. For a more comprehensive list of the model assumptions, see Appendix A.

b. Simkit Model Assumptions

The first assumption is that the Simkit replica is a good distillation of the MODSIM model and that it correctly imitates the model as it stands. The MODSIM code is not accessible due to its proprietary nature. Therefore, we ensure that a similar arrivals process is achieved while producing similar output in all other areas. Another assumption of the model is that approximations made about past weather conditions are good predictors for future events. Also, that the efforts placed into approximating weather are sufficient for Eglin AFB.

F. LITERATURE REVIEW

Mustafa Azimetli's (2008) thesis, "Simulation of Flight Operations and Pilot Duties in LANTIRN Fighter Squadrons using Simkit," uses a weather simulation to determine how often they can keep their Low-Altitude Navigation and Targeting Infra-Red for Night LANTIRN fighter squadrons up and running. His analysis uses a weather simulation that splits the weather into three categories of flight. The three categories allow certain pilots to fly, depending upon their training levels, etc. He uses a Markov chain process to create current weather categories. Similarly, this thesis uses a Markov chain process to generate weather. However, this thesis uses probabilities to go from good to bad weather, and vice versa, moving students in and out of a weather wait queue. All pilots are in a training environment and the breakup of the experience categories is not a necessary piece. Azimetli uses METAR reports that are simulated from weather reports in Turkey. Although he uses a similar weather simulation, his simulation evaluates the use of the LANTIRN A/C. This thesis gives more realistic estimates to the current TTT estimates given by the ITC model.

Axtell, Axelrod, Epstein, and Cohen (1996) coin the term docking, for making two simulation models match. In their case study, they compared two different models to see if they could accomplish the same task. Similarly, this thesis compares two different

models—scripted in different software, by different authors, and at a different time—to accomplish the same thing. The ability to compare models is an important component of this thesis; however, it differs in its attempts as it does not replicate the entire model, but only a small part of it. Docking as a thread of research is primarily found in the agent-based modeling community and, to our knowledge, this is the first application to discrete event simulation.

This thesis develops a Simkit version of the flight event process and docks it to the ITC model. Using Azimleti's thesis as a point of departure, we modify the Simkit model to more appropriately model the effects of weather on flight operations.

G. THESIS OUTLINE

Chapter II discusses the ITC model makeup. It also discusses the analysis that shows how important weather is to this model. Chapter III is the Simkit replica of the A/C flight event found in the ITC model. We discuss how it runs and show the analysis of how it functions. Chapter IV reviews the proposed changes to the weather model using Simkit. Analysis is done to illustrate how weather changes affect resource utilization in this improved Simkit model. It also does a comparison analysis of the two different Simkit models showing the significance of the proposed changes. Chapter V provides the conclusions, recommendations, and suggestions for further research using this model.

THIS PAGE INTENTIONALLY LEFT BLANK

II. ITC MODEL AND WEATHER

A. ITC MODEL

The ITC model covers the entire life cycle of a student pilot, from inception to graduation. It takes the pilots through a set of events, controlled by syllabi, using MODSIM III, which is a DES modeling software program that controls the flow of students through the training cycle (Kenny, 2010). The time it takes to go through this training process is TTT.

As explained in Chapter I, there are many inputs and outputs. The inputs are described as variables and can be found in Table 1. In the ITC model, all of the variables found in Table 1 can be manipulated in the simulation. However, most of them are prewritten into a COA worksheet that is then read into the model using text files. These text files create the scenario and control a number of factors running the ITC model. Some of these text files are the syllabus schedules that pilots go through. Others are the plan for purchasing A/C and utilization of IPs. All of the 36 variables are written into the text files and read into the ITC model at the instantiation of the run. The GUI screens, depicted in Appendix B, allow the user to access the individual variables. In this way, changes can be made to the variables without touching the individual text files. There are a lot of output measures that can be analyzed. In this chapter, we focus upon the most significant of these outputs, mean TTT.

Table 1. All of the decision and noise variables controllable by the GUI.

Decision Variables	Noise Variables
Unclassified Classrooms	Winter Cancellation Rates
Classified Classrooms	Summer Cancellation Rates
Interactive Courseware Work Stations	Probability of Mechanical Failure
FMSs	Time for Repairs
Mission Readiness Trainer (MRT) Flight Simulator	A/C Flight Preparation Time
Mission Planning Station	A/C Refly Rate
Brief/Debrief Rooms	Simulator Refly Rate
A/C Two-Seat	FMS Planning Time
A/C Systems Maintenance Trainer	FMS Briefing Time
Weapons Load Trainer	FMS Debriefing Time
Ejection System Maintenance Trainer	MRT Planning Time
Engine Trainer	MRT Briefing Time
Outer Mold Line Train Lab	MRT Debriefing Time
Military Operating Area Slots	
Low Levels	
Target Slots	
Runways	
IPs	
Contract Instructors	
Training Days/Year	
Hours in a Day	
Choice of Syllabus	
Induction of IPs	

The ITC model estimates the amount of time pilots take to go through the training cycle, while encountering a number of possible delays. As this model is written, pilots must advance from one training event to the next in order—no skipping or changing of events is allowed. If delays occur, they occur at the time the event is scheduled. Weather is just one of the delays that can occur during the training cycle. This thesis will show that weather is an important factor in this model.

B. ITC WEATHER

The A/C flight event is one of many events simulated in the ITC model. The flight event is read in as part of the syllabus found on the text files. The flight events are the basis of this analysis, as it is the only location where weather affects the TTT

estimates. Each syllabus has a certain number of hours that it requires pilots to fly. The flowchart for the flight event and weather is found in Figure 1. The flowchart shows the process that a student goes through in order to fly a scheduled event. It also shows the approximate times that it takes in order to go through the briefs and possible maintenance cycle.

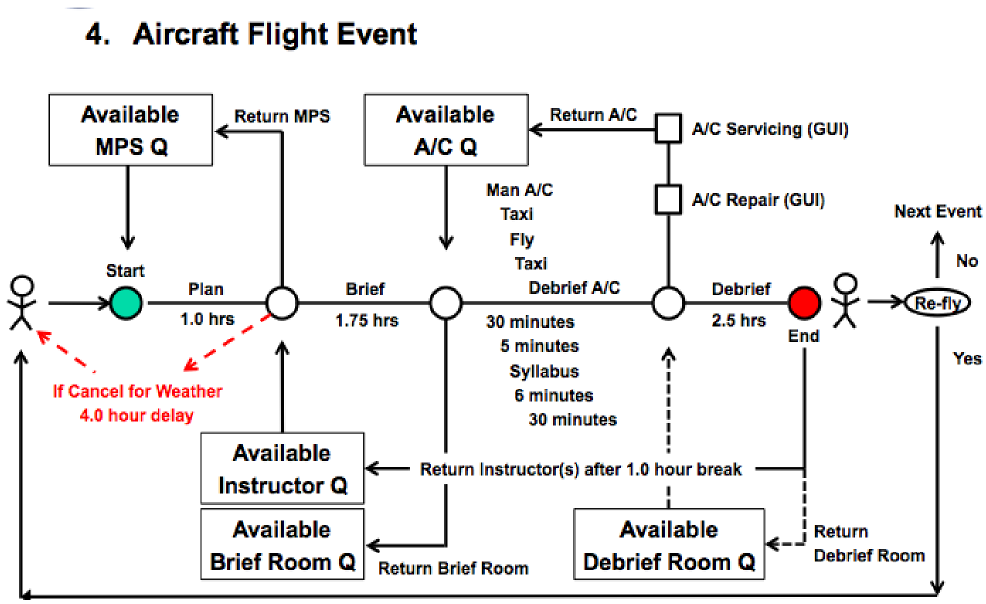


Figure 1. The flowchart on which the ITC model bases its A/C flight events.

Weather is a very simple component of the ITC model. It uses a system that divides the year into two different seasons—summer and winter. Both the summer and winter cancellation rates can be adjusted using the Scenario User interface found in Appendix B. The effect of weather is simulated using a Bernoulli trial. A uniform random number is drawn and compared to the designated cancellation rate. If the number is greater than the cancelation rate percentage during that season, then the A/C is allowed to fly. Those that fall under the rate are then placed in the weather wait queue for four hours while the system continues on. As already discussed, other students coming into the queue directly after the one that is cancelled are checked with a similar Bernoulli trial and will either fly or go into a weather wait queue independent of what happened to the student prior.

C. DESIGN OF EXPERIMENTS (DOE)

1. Measure of Effectiveness (MOE) and Factor Selection

An MOE is "a qualitative or quantitative measure of the performance of a model or simulation or a characteristic that indicates the degree to which it performs the task or meets an operational objective or requirement under specified conditions" (MOE, 1998, p. 136). The MOE used in this section is mean TTT.

The variables are usually divided into three groups: decision factors, noise factors, and artificial factors. The decision factors are those that are controllable in the real world; the noise factors are uncontrollable real-world factors; and the artificial factors are those specific to simulation, such as initial state and warm-up periods (Sanchez, 2000). Table 1 is a description of the variables and which ones are used in the DOE for this analysis. The key variables highlighted in Table 1 are summer and winter cancelation rates, A/C and simulator refly rates, probability of mechanical failure rates, time to repair A/C, and A/C preparation time. The weather factors are vital to the analysis in this thesis. The others are chosen because planning does not control their outcome. All the decision variables are input automatically according to the latest COA, and, although they affect the overall TTT, keeping them constant allows one to analyze the most important data in the uncontrollable factors. By adjusting the noise variables, we are better able to understand how much effect uncontrollable variables will have on the overall TTT. This knowledge helps decision makers create better training schedules by allotting an appropriate amount of time for the training. It also gives them the ability to make a case for necessary resources.

2. Nearly Orthogonal Latin Hypercubes (NOLHs)

An initial design of seven factors makes it impossible to run a model with the entire set of possible parameter space. A full-factorial design, with 10 factors and 10 levels each, would require 10 billion design points. In an example using 20 replications, 200 billion runs of simulation would be required. In the event that each simulation took one second, the entire experiment would take over 6,300 years to run. A much more efficient design is needed (Sanchez, 2008).

To construct a more efficient design, we use an NOLH experimental design spreadsheet developed by Dr. Susan M. Sanchez at the Naval Postgraduate School (NPS) in 2005. A NOLH design facilitates an investigation into the domain of numerous factors. These designs allow one to see the effects of one factor independent of other factors. They also maintain the benefits of orthogonality, such as independence of estimates of variables, while providing the added benefit of space-filling properties giving superior analytical performance. This benefit allows one to explore a model with more depth. In this thesis, we evaluate seven factors, as shown in Table 2. To cover more of the design space, the $(N_0)_{11}^{33}$ sheet is used with a wrapping technique (Cioppa, 2002). This gives a more appropriate response surface to test for nonlinear behavior. It is important to note that these designs lose very little in orthogonality, which means that we stay away from multicollinearity issues (Cioppa & Lucas, 2007).

Table 2. NOLH design spreadsheet with factor names and levels (9 of 65).

low level	0	0	0	1	0	0	0
high level	60	60	99	24	40	40	4
decimals	0	0	0	1	0	0	0
factor name	WxCanc	SxCanc	ProbA/Cfail	TimeToRep	FightRefly	SimRefly	NextSortie
	60	6	43	5.3	35	25	3
	54	60	12	9.6	19	8	3
	53	26	90	4.6	1	24	3
	34	53	99	10.3	38	6	3
	56	2	46	6	28	29	2
	58	56	31	7.5	18	9	1
	41	28	96	6.8	0	26	2
	32	41	93	8.9	36	10	1
	39	15	22	13.2	29	13	0

Using the $(N_0)_{11}^{33}$ sheet and wrapping it once results in 66 design points. In order to ensure independence, the center design point is deleted from the second group of numbers, leaving a total of 65 design points. These 65 design points are then run through the ITC model with 50 replications each, totaling 3,250 runs of simulation, with

120 pilots going through each run, giving information on 390,000 pilots in 14 different syllabi. Of the 14 syllabi, two are upgrades and the TTT levels are significantly shorter.

The space-filling properties of the DOE are shown in Figure 2. The points in the matrix are distributed evenly across the matrix. This strengthens the design and allows access to 2nd order effects.

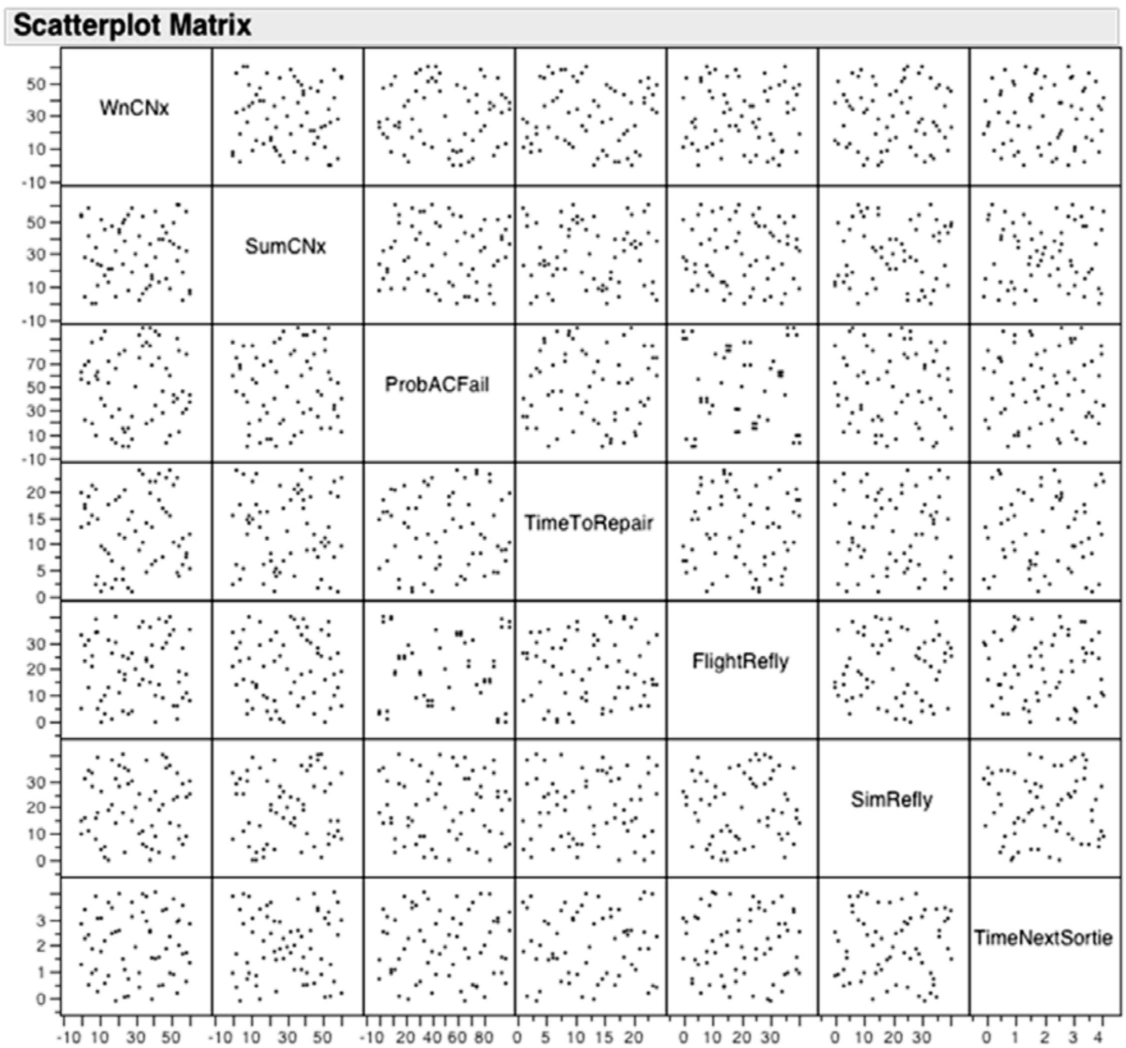


Figure 2. The space-filling properties of the 7-factor DOE using a scatterplot matrix.

D. DATA ANALYSIS

Before a full analysis on the data is performed, it is necessary to do preliminary work on the output information. First, we take into account the warm-up and cool-down

periods of the ITC model. To account for these periods, 150 days is chopped from the beginning and end of the data. Those points that fall in the first 150 days and end in the last 150 days are not used in the analysis. Next, the output of the run is broken down into many categories. An example of the output can be found in Table 3. The first column describes which pilot (1-120) was training and the next column gives a class number, depending upon the start date. The third column is the syllabus each pilot flew, the fourth is the time they started that syllabus, and the fifth is the completion date. Next, it gives the overall TTT; followed by calendar weeks; the approximate number of months, weeks, and days according to the syllabus start time; and scheduled no training days. Then, it gives the study planning time, followed by flight hours and simulation hours of use in the syllabus.

Table 3. Output spreadsheet from simulation runs (11 of 390,000 rows shown).

Pilot	Class Number	Class.Type	Start Day	Start Time	Finish Day	Time to Train	Calendar Weeks	MO/WK/ DAYS	Study Planning Hours Day	Flight Hours	Sim Hours	CQ Refly	DP	rep
Pilot0001	1	DCMA_Initial_Cadre.txt	411	7:00:00	460	50	11	2:02:01	2	8	29	NA	0	1
Pilot0002	NA	DCMA_Initial_Cadre.txt	411	7:00:00	462	52	12	2:02:03	2	8	34	NA	0	1
Pilot0005	3	DCMA_Initial_Cadre.txt	461	7:00:00	498	38	9	1:03:04	0	8	36	NA	0	1
Pilot0006	NA	DCMA_Initial_Cadre.txt	461	7:00:00	502	42	9	2:00:03	0	10	31	NA	0	1
Pilot0003	2	B1.0_TX_ALL_B_Eglin.txt	431	7:00:00	507	77	17	3:03:04	0	24	49	NA	0	1
Pilot0004	NA	B1.0_TX_ALL_B_Eglin.txt	431	7:00:00	507	77	17	3:03:04	0	25	53	NA	0	1
Pilot0012	NA	B1.0_TX3_A_Eglin.txt	486	7:00:00	531	46	10	2:01:02	1	12	19	NA	0	1
Pilot0011	6	B1.0_TX3_A_Eglin.txt	486	7:00:00	531	46	10	2:01:02	1	12	21	NA	0	1
Pilot0014	NA	B1.0_TX4_A_Eglin.txt	486	7:00:00	538	53	12	2:02:04	1	18	25	NA	0	1
Pilot0013	7	B1.0_TX4_A_Eglin.txt	486	7:00:00	538	53	12	2:02:04	0	15	22	NA	0	1
Pilot0008	NA	TX3_B1.0.txt	461	7:00:00	574	114	26	5:03:02	1	36	62	NA	0	1

The information found in the output is specified by many different class types, as seen in Table 3. After some simple analysis of all of the different classes, we are able to show that each class performs similarly. This thesis uses a single class to describe weather and its effects. This was to limit the repetition in graphs and charts. In order to choose a single class, we had to analyze which class would be the most appropriate to generalize all classes over. Of the 14 syllabi, the least similar syllabi are the upgrades, the initial cadre group, and the other three syllabi that lead into the two upgraded syllabi. Due to their shortened length of training days, they would not be helpful in predicting overall TTT for a student going through a normal training cycle.

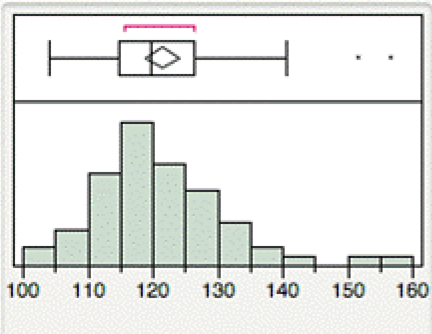
Table 4 shows summary statistics for all 14 of the different classes based upon TTT; the first six are the syllabi explained above. That left eight syllabi to decide from. Being as they are all similar to the others, we choose one with an average length of training days and the largest number of students going through it. The syllabus TX3_B2.0.txt has 58,500 pilots train using its syllabus under differing design points and replications.

Table 4. Summary statistics on the 14 different syllabi. Highlighted is the class used in this analysis.

	Class.Type	N Rows	Median (Training.Day	Mean(Training.Days)	Std Dev(Training.Days)	Min (Training.D	Max (Training.D
1	B1.0_TX_ALL_B_Eglin.txt	19500	63	64.29374359	6.829724182	49	96
2	B1.0_TX3_A_Eglin.txt	6500	42	42.498461538	3.946616085	34	66
3	B1.0_TX4_A_Eglin.txt	6500	50	50.553692308	4.7101600004	41	74
4	B2.0_TX3_Upgrade.txt	78000	20	21.331846154	4.7666835423	15	66
5	B2.0_TX4_Upgrade.txt	32500	22	23.034769231	5.529115333	16	71
6	DCMA_Initial_Cadre.txt	13000	48	44.728769231	6.4425577031	32	63
7	TX1_B1.0.txt	6500	118	121.12861538	17.608348245	90	197
8	TX1_B2.0.txt	6500	136	138.71692308	14.840440472	109	200
9	TX3_B1.0.txt	45500	104	105.37169231	11.280036999	81	185
10	TX3_B2.0_Env.txt	19500	138	139.0865641	11.718394921	110	183
11	TX3_B2.0.txt	58500	119	121.42523077	12.392685696	96	201
12	TX4_B1.0.txt	39000	111	113.02441026	12.507590722	85	189
13	TX4_B2.0_Env.txt	26000	148	149.19484615	12.414949178	120	191
14	TX4_B2.0.txt	32500	126	127.52393846	12.32879603	101	212

To better understand the response variable, we show a histogram with some of the summary statistics in Figure 3. This shows that the mean TTT is about 121 days to train and the standard deviation is about 10 days. The response variable is somewhat symmetric with what seem to be a couple of outliers. The outliers seem to be caused by a high rate of A/C failure, as well as an extreme amount of time to repair.

Mean(Training.Days)



Quantiles

100.0%	maximum	156.756
99.5%		156.756
97.5%		153.498
90.0%		134.391
75.0%	quartile	126.259
50.0%	median	119.836
25.0%	quartile	114.888
10.0%		110.461
2.5%		104.416
0.5%		103.967
0.0%	minimum	103.967

Moments

Mean	121.42523
Std Dev	10.141568
Std Err Mean	1.2579067
Upper 95% Mean	123.93819
Lower 95% Mean	118.91227
N	65

Figure 3. Distribution of the response variable.

An ordinary least squares (OLS) model is fit to mean TTT with the seven factors, to include all second-order interactions. Stepwise corrected Akaike Information Criterion (AICc) forward helps determine those factors that are important (SAS, 2009). Figure 4 shows the progression of R^2 for the AICc stepwise model selection process. The vertical line shown on the graph depicts the best fit model according to the AICc. The adjusted R^2 is about 0.90, as shown in Figure 5. This does not mean that we have the best fit model. In fact, it may indicate that we over fit the model. Using the AICc method, we kept 14 of the original 28 significant terms. More analysis is vital in order to decipher if we have chosen a good model.

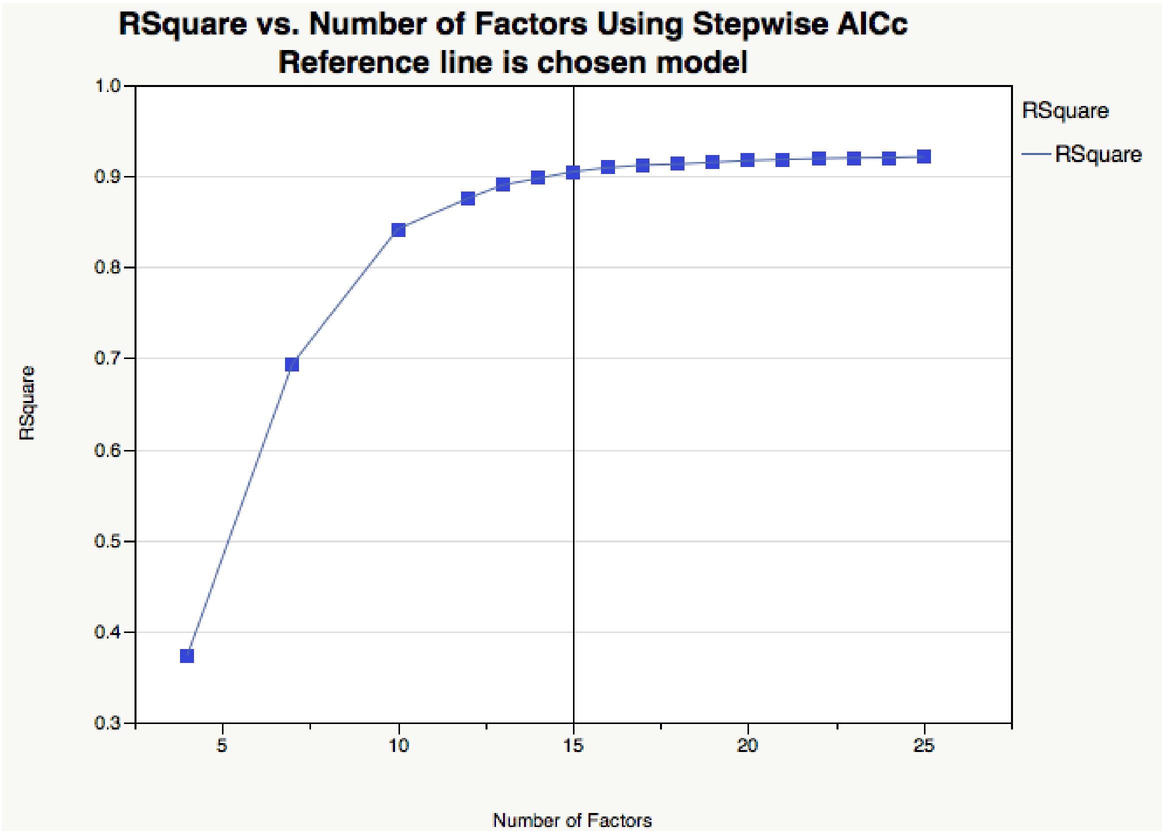


Figure 4. R^2 progression with number of significant factors.

Summary of Fit	
RSquare	0.904938
RSquare Adj	0.87832
Root Mean Square Error	3.537647
Mean of Response	121.4252
Observations (or Sum Wgts)	65

Figure 5. Summary of fit on the linear model using mean TTT as the MOE.

The sorted parameter estimates help one to identify the important factors in the model and how they rank accordingly. Summer and winter cancellation rates are both among the most highly statistically significant factors in the model. The OLS model, shown in Figure 6, illustrates that both summer and winter cancellation rates are important in the model. Every five percentage-point change in either summer or winter

cancellation rates equates to approximately a one-day increase in mean TTT. For example, increasing the summer cancellation rate from 0.2 to 0.25 increases expected mean TTT by approximately one day.

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
FlightRefly	0.3417278	0.036547	9.35	<.0001*
SumCNx	0.2193404	0.024438	8.98	<.0001*
ProbACFail	0.1156315	0.014759	7.83	<.0001*
TimeToRepair	0.4822408	0.063631	7.58	<.0001*
WnCNx	0.1754402	0.024436	7.18	<.0001*
(ProbACFail-49.5077)*(TimeToRepair-12.5031)	0.0189804	0.002653	7.15	<.0001*
SimRefly	0.1855317	0.036547	5.08	<.0001*
(ProbACFail-49.5077)*(TimeNextSortie-2.01231)	0.0392162	0.013148	2.98	0.0044*
TimeNextSortie	1.0805913	0.365444	2.96	0.0047*

Figure 6. Sorted parameter estimates using mean TTT as the MOE.

E. VALIDATION

It is important to evaluate the residuals versus the predicted plot in order to validate the model. Figure 7 depicts the predicted mean training days' values versus the residuals. The points seem to have no particular pattern, demonstrating constant variance. This supports the first assumption that there is likely constant variance.

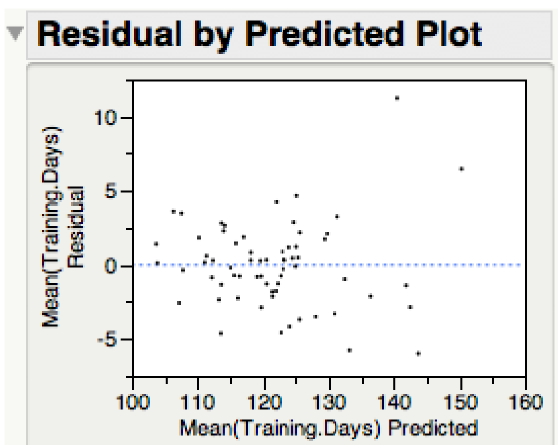
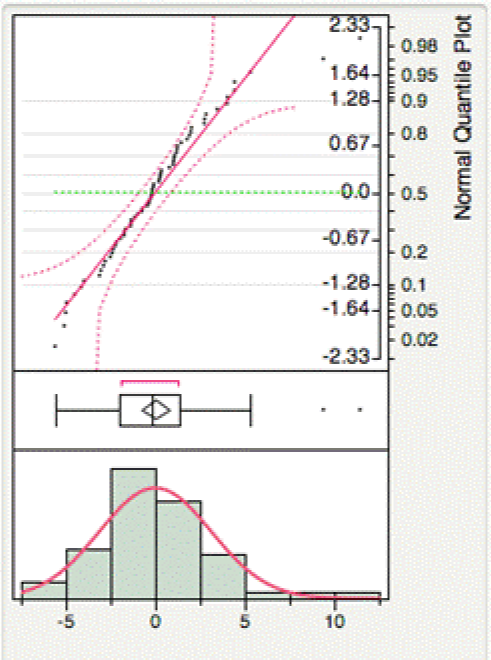


Figure 7. Mean TTT residuals by predicted plot.

According to the qq Normal plot, we can see that there is some granularity, as well as being a little right skewed, as shown by the bar chart in Figure 8. Overall, it seems to show that the points are approximately normally distributed.

Residual Mean(Training.Days)



Quantiles

100.0%	maximum	11.4364
99.5%		11.4364
97.5%		10.1032
90.0%		4.01763
75.0%	quartile	1.34037
50.0%	median	-0.1812
25.0%	quartile	-2.0567
10.0%		-4.0582
2.5%		-5.2715
0.5%		-5.605
0.0%	minimum	-5.605

Moments

Mean	4.373e-15
Std Dev	3.1268676
Std Err Mean	0.3878402
Upper 95% Mean	0.7747998
Lower 95% Mean	-0.7748
N	65
Sum Wgt	65
Sum	2.842e-13
Variance	9.7773008
Skewness	1.0227519
Kurtosis	2.550032
CV	7.151e+16
N Missing	0

Figure 8. Normal QQ plot of residuals with histogram.

F. CONCLUSIONS

Weather is important in the ITC model as currently modeled. Using linear regression analysis, we show that winter and summer cancellation rates are significant factors. We use a model that maximizes adjusted R^2 without over-fitting the model. Through the use of residual analysis, the assumptions of normality are tested. We conclude that the model is a valid model. Everything found supports the claim that weather is important in the ITC model. All analysis in this thesis is demonstrated through the exclusive use of JMP Statistical Discovery software.[1]

[1] Additional information about the JMP Statistical Discovery software can be found at www.jmp.com.

III. REPLICA SIMKIT MODEL

A. SIMKIT MODEL DESIGN

The purpose of this thesis project is to examine the manner in which weather is modeled in the ITC model, to propose improvements, and to quantify the benefits. The most effective way to accomplish this is to modify the ITC model directly, and compare the legacy and improved models. However, the ITC model cannot be modified for legal and financial reasons. A replica of the ITC model flight event management process is created and implemented in Simkit. This is the important part, because it is the only place weather has an effect on TTT.

A flow chart of the Simkit replica model that recreates the A/C flight events found in the ITC model is shown in Figure 9. The flow in the ITC model and the replica model are similar. The primary difference between the replica model and the ITC model is that the replica model has unconstrained resources. For example, the model is not limited to a certain number of A/C, brief rooms, or IPs.

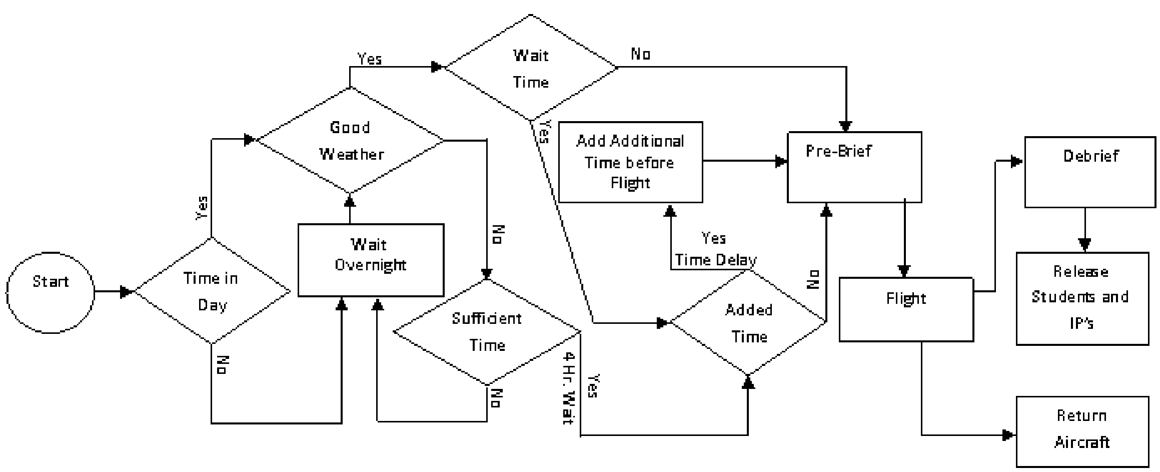


Figure 9. Flowchart of the Simkit replica model.

The Simkit replica model is implemented in 10 different Java classes. An event graph diagram of this model is found in Appendix C. The following is a brief description of how the model works. First, an arrival process governs the arrival of flights into the system. In the current implementation, the interarrival times are exponentially

distributed, which is an appropriate assumption given the necessity to abstract away from all nonflight-related evolutions on the syllabi. One of the parameters over which we data farm is the shape parameter.

Next is the Flight and FlightCreator process. As flights arrive according to the arrival process, the Flight Creator class creates Flight objects. A flight object possesses the number of students involved in the flight, the length of the flight, as well as the time at which the flight object entered the system. The number of students required for each flight is either one or two. Similarly, the flights on a given syllabus, for a particular course of study, may range from approximately one to three hours. In order to maintain a sense of generality, the FlightCreator determines the number of students and duration for a flight according to probability distributions. A simple Bernoulli trial is used to determine the number of students, while a Triangular distribution, with a lower bound of 0.5 and an upper bound of 3.0, is used to determine duration. The probability of the Bernoulli trial and the mode of the Triangular distribution are varied according to the experimental design.

If there is sufficient daylight remaining to fly an event, the system checks the weather according to a Bernoulli trial. The probability of the Bernoulli trial depends upon the season. The SeasonChanger class simply alternates between summer and winter, thereby enabling the WaitCollector class to utilize the appropriate weather cancellation rate. If the random number it generates falls below the assigned cancellation rate for that period, flights are sent into a weather wait queue.

In reality, a flight that is delayed due to weather may find that the required training range is no longer available. Similarly, in the ITC model, a pilot may emerge from a weather delay and have to wait for an IP to become available. In order to account for this phenomenon in the replica model, every flight that emerges from the weather delay queue experiences a Bernoulli trial to determine if an additional delay will be applied. If so, an additional delay of from 1 to 5 days is imposed in order to simulate the possibility of delays exacerbated by the weather delay. This additional wait time

accounts for some of the time delays associated with the unavailability of resources. Students, for whom additional flight wait time is imposed, do not undergo another weather check.

The BriefClass processes the remainder of the flight event. It manages the utilization of resources such as brief rooms, IPs, A/C in use, and completed flights. At various points in the flight evolution, the brief rooms, IPs, and A/C are released. Finally, the ITCModelExe class instantiates all the relevant objects and manages the entire process. This class listens to all the other processes while recording the output information.

Other important features in this model are the FIleIO, ReadCommaSeperatedInput, and the OutputCatcher processes. The first two read in Comma Separated Value (.csv) sheets and convert them into arrays that are used for farming later. The first sheet, DOE.csv, is the design spreadsheet that has all the factors and their assigned values. The second sheet is Seed.csv, which is a seeding worksheet. This worksheet provides the ability to replicate the experiment elsewhere, using a different machine, and still get the same values. The OutputCatcher process allows access to the output. It captures the data on a separate screen and saves it to a comma separated text file. This allows JMP to access the information for processing.

B. SCOPE, LIMITATIONS, AND ASSUMPTIONS

It is important to replicate the original model closely, in order to obtain useful results. This model does not recreate the entire ITC model. It recreates the flight process portion of the model only. This is the only place where weather has an effect upon the training process found in the ITC model. We compare the number of arrivals from the ITC model to the number of arrivals coming through this model, and we find that the arrival process is similar to the ITC model's arrival process.

Resources are not restricted in this model. In order to overcome this weakness, an additional delay is placed upon some of the flights that are delayed due to weather. This rescheduling delay is simulated by an additional delay time of up to five days, as explained in the WaitCollector process.

C. EXPERIMENTAL DESIGN

In the interest of time and simplicity, the model assumes unlimited resources such as briefing rooms, IPs, and A/C required. All resources are measured, which helps decision makers understand resource usages. Resource utilization rates are analyzed as well, to show differences between the replica model and the model found in Chapter IV.

The MOE for this model is slightly different than the MOE for the ITC model. This model does not estimate total mean TTT because the model only includes flight events. Instead of mean TTT, it uses Mean_DWQ. DWQ is a response variable found in Table 6; it is the mean delay in wait queue by design point.

The list of decision variables is also different from the Simkit model. This model does not limit the resources, so many of the decision variables are output instead of input. Table 5 is a list of decision and noise variables found in the replica model.

Table 5. List of Decision and Noise Variables in the Simkit replica model.

Decision Variables	Noise Variables
Arrival Rates	Winter Cancellation Rates
Number of Students per Flight	Summer Cancellation Rates
Flight Duration	In-Brief Times
IP Break Time	Debrief Time
	Aircraft Flight Preparation Time
	Post Flight Preparation Time
	Maintenance Time
	Additional Time

This thesis implements the 12 factors listed in Table 5 in a $(N_0)_{16}^{65}$ NOLH design. It replicates each design point 50 times, for a total of 3,250 runs. We use common random numbers for each replication, so as to reduce variance due to random number selection between runs (Law & Kelton, 2000). The simulation completes this experiment in approximately 15 minutes on a desktop computer, with a Pentium 4 central processing unit, 3.6 gigahertz speed, and 3.0 gigabytes of random access memory. Common random

numbers are used to decrease the variance caused by simulation (Law & Kelton, 2000). Appendix D is a scatterplot matrix that shows the space-filling properties of our design space.

D. ANALYSIS

The model exports its output data in a comma separated text file, which facilitates import into JMP. A sample of the output is shown in Table 6.

Response variables include DWQ, mean and maximum number of students in the weather wait queue, as well as the mean and maximum number of students in the wait queue. Additional response variables include max A/C in use, maximum IPs in use, and maximum brief rooms in use.

The simulation runs for a three-year period, similar to the ITC model. This model is not limited by resources, but it is able to show resource usage. It also accounts for resources, to give decision makers the ability to identify resource utilization rates.

Table 6. A part of the output file that is saved as a .csv file.

rep	dp	time	wxWaitQ_mean	wxWaitQ_max	DWQ	wxWaitQ_delay_max	wxWaitPersQ_mean
0	0	36000	0.67	8	8.9	144	1.31
0	1	36000	0.89	10	9.84	149	1.12
0	2	36000	0.73	8	8.56	144	1.2
0	3	36000	0.53	7	7.47	144	0.6
0	4	36000	0.13	7	1.5	138.42	0.21
0	5	36000	0.66	10	9.99	147.62	0.72
0	6	36000	0.29	6	3.76	138.83	0.52

A histogram and summary statistics of DWQ are shown in Figure 10. The mean response is 14.3 hours, with a standard deviation of 7.2 hours. All mean delay times fall at or below 33 hours.

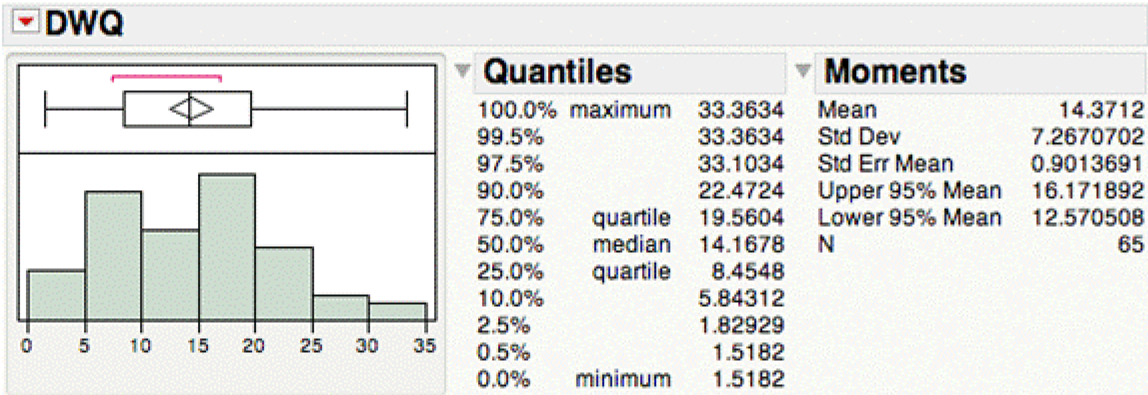


Figure 10. Distribution of the response variable Mean_DWQ.

A partition tree is used to explore the data and identify important factors. The first four splits of the tree are all weather-related variables, WxCanc and SxCanc, which confirms their relative importance. Adjusting weather at different levels allows one to account for the majority of the variation in the response variable. These five splits (shown in Figure 11) result in an R^2 of approximately 0.8.

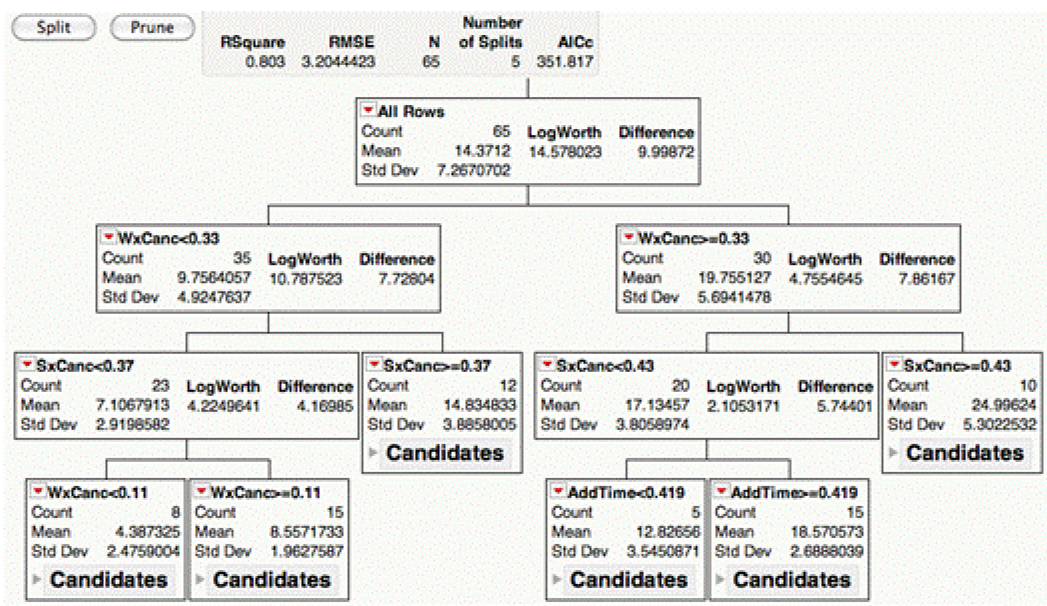


Figure 11. Partition tree with Mean_DWQ as a response variable.

The partition tree also keeps track of R^2 at every split and Figure 12 is a graphical representation of the R^2 at every split. As seen below, very little predictive power is

gained by going beyond the 5th split. This helps to determine a stopping point. The steps in Figure 11 only show up to five splits, whereas Figure 13 gives the history out to nine steps.

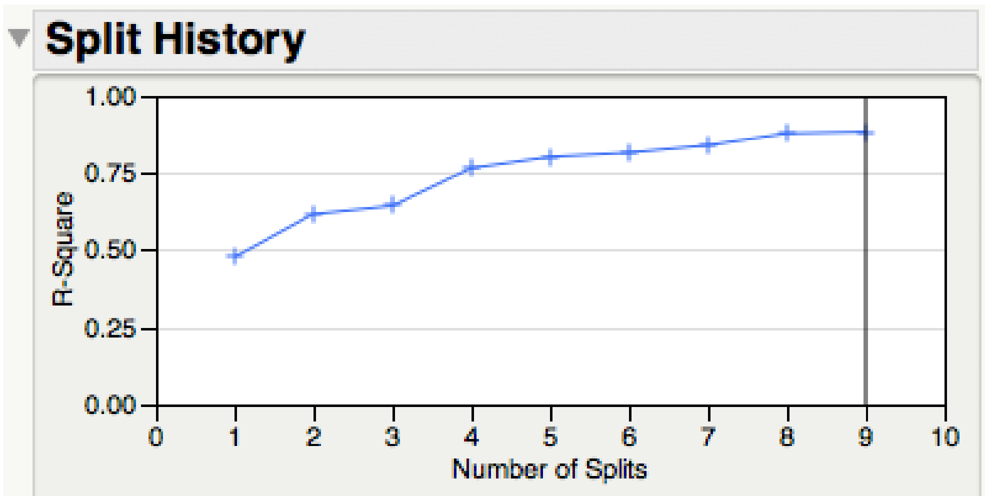
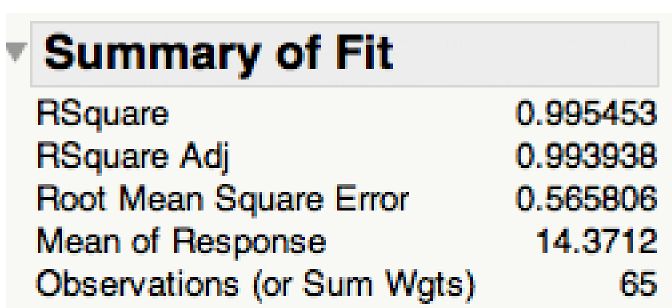


Figure 12. Partition tree split history.

An OLS model is fit to Mean_DWQ with all 12 factors, to include second-order interactions. With Mean_DWQ as the MOE and using the AICc method, 17 of the original 79 factors enter the model, which results in an adjusted R^2 of 0.993, as shown in Figure 13.



Summary of Fit

RSquare	0.995453
RSquare Adj	0.993938
Root Mean Square Error	0.565806
Mean of Response	14.3712
Observations (or Sum Wgts)	65

Figure 13. Summary of fit for OLS model using Mean_DWQ as MOE.

The sorted parameter estimates show how important the factors are in this model. Analyzing the sorted parameter estimates shown in Figure 14 reveals that SxCanc and WxCanc rates affect Mean_DWQ significantly. Increasing weather rates from, say,

0.3–0.35, increases the mean delay time by about 1.5 hours. The AddTime parameter is the probability that a flight experiences a delay, in addition to a weather delay. The sign is positive, as expected. For every 0.1 increase in the probability of added time, approximately two additional hours are spent in the weather wait queue. In addition, it is unsurprising that WxCanc*AddTime and SxCanc*AddTime are the two most important interactions. As the weather cancellation rates increase, the addTime becomes more important. The weather delays are expected to be important, but arrival rates (Lambda) do not make as significant an impact as expected.

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
WxCanc	30.59688	0.399757	76.54	<.0001*
SxCanc	23.400662	0.399766	58.54	<.0001*
AddTime	18.507409	0.598663	30.91	<.0001*
(WxCanc-0.30031)*(AddTime-0.50012)	33.649145	3.398631	9.90	<.0001*
(SxCanc-0.30031)*(AddTime-0.50012)	26.75647	3.137364	8.53	<.0001*
(%Studs-0.50031)*(WxCanc-0.30031)	9.5485706	1.727843	5.53	<.0001*
(Lambda-0.06502)*(InBrf-1.75031)	-73.73404	15.12965	-4.87	<.0001*
(FltDur-1.75077)*(Post-0.50015)	3.046066	0.712945	4.27	<.0001*
(Lambda-0.06502)*(Maint-1.00154)	-74.30098	17.99443	-4.13	0.0001*
FltDur	0.289362	0.095802	3.02	0.0040*
(WxCanc-0.30031)*(SxCanc-0.30031)	-3.196259	2.416072	-1.32	0.1921
Post	0.2313033	0.479097	0.48	0.6314
Maint	0.037149	0.236516	0.16	0.8759
%Studs	0.0250762	0.239604	0.10	0.9171
InBrf	0.0188482	0.239619	0.08	0.9376
Lambda	0.0401288	4.790817	0.01	0.9934

Figure 14. Sorted parameter estimates of the OLS model using Mean_DWQ as MOE.

E. VALIDATION

1. Regression Validation

Figure 16 shows the residuals by predicted plot, which demonstrates constant variance. The correlation of estimates coefficients matrix shows that no correlation above 0.17 exists between any of the terms, which indicates there are no issues with multicolinearity. These estimates fall well below 0.9, which is the level Hamilton (1992) discusses as high correlation. A qq-plot of the residuals is shown in Figure 15, as is the

residual by predicted plot and goodness-of-fit test. All of these indicate that the errors are sufficiently distributed in a normal manner. This model satisfies the necessary assumptions for an OLS model.

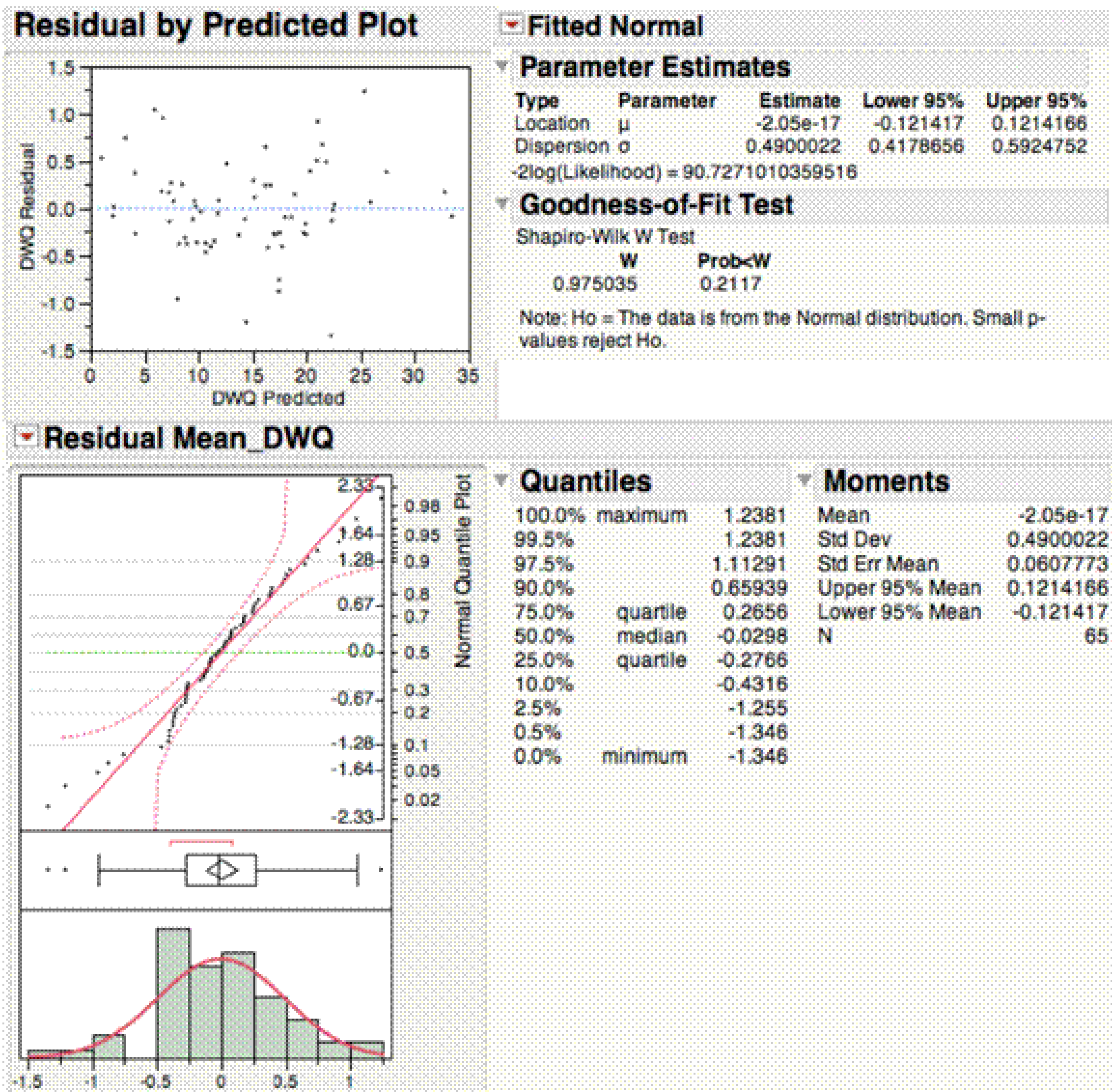


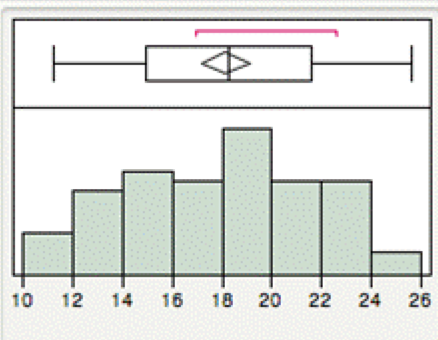
Figure 15. Mean_DWQ Residuals by Predicted Plot, QQ normal plot of residuals with histogram and goodness-of-fit test.

2. Maximum Aircraft Utilization

One of this model's important simplifying assumptions is that resources such as briefing rooms, IPs, and A/C in use are unconstrained. In order to evaluate the impact of

these assumptions, it is necessary to examine the areas of the design space where any of these resources exceed their maximum capacities. This section examines Max_A/CInUse, which is the mean of the maximum number of A/C flown by design point. A fully operational training squadron is only assigned 20 A/C and only a portion of those are expected to fly on any given day. Therefore, it is reasonable to conclude that for areas of the design space where Max_A/CInUse exceeds 20, the estimate of Mean_DWQ is a lower bound. This is because in the ITC model, as in reality, flights that are delayed due to weather may subsequently be delayed because the aircraft originally intended for the flight is reassigned to a different flight. Figure 16 gives the distribution of the Max_A/CInUse for the 65 design points. Appendix E shows the distributions of the Max A/C, Max brief rooms, and the Max IPs in use over the 3,250 data points.

Max_A/CInUse



Quantiles

100.0%	maximum	25.64
99.5%		25.64
97.5%		24.808
90.0%		23.448
75.0%	quartile	21.58
50.0%	median	18.24
25.0%	quartile	14.94
10.0%		13.128
2.5%		11.396
0.5%		11.24
0.0%	minimum	11.24

Moments

Mean	18.147077
Std Dev	3.7751299
Std Err Mean	0.4682472
Upper 95% Mean	19.082508
Lower 95% Mean	17.211646
N	65

Figure 16. Distribution of the Max_A/CInUse by 65 design points.

These distributions acknowledge how each resource is utilized and whether there is enough of that resource available. If there are insufficient resources available, then longer delays will occur. The additional time added in the model attempts to account for the lack of resource restriction found in this model.

An OLS model is fit to Max_A/CInUse with all 12 factors, to include all two-way interactions. Stepwise AICc is employed to ascertain the appropriate factors. This results in an adjusted R^2 of 0.984 (shown in Figure 17). The sorted parameter estimates are also shown in Figure 17. The most important parameter is %Studs. As the proportion of flights with two students increases, i.e., as %Studs decreases, Max A/C In

Use increases. With A/C usage as our parameter, it makes sense that the number of students flying, and how often they are flying, is near the top of the predictive parameters.

Summary of Fit

RSquare	0.988627
RSquare Adj	0.984836
Root Mean Square Error	0.464883
Mean of Response	18.14708
Observations (or Sum Wgts)	65

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
%Studs	-10.6671	0.196853	-54.19	<.0001*
Lambda	131.10362	3.936608	33.30	<.0001*
WxCanc	-1.999293	0.328478	-6.09	<.0001*
SxCanc	-1.961494	0.328461	-5.97	<.0001*
(%Studs-0.50031)*(WxCanc-0.30031)	9.4800284	1.718393	5.52	<.0001*
(SxCanc-0.30031)*(Maint-1.00154)	11.587136	2.788456	4.16	0.0001*
(InBrf-1.75031)*(Maint-1.00154)	3.0552142	0.784305	3.90	0.0003*
FltDur	0.3022131	0.07871	3.84	0.0004*
(%Studs-0.50031)*(InBrf-1.75031)	-4.577881	1.378481	-3.32	0.0017*
(WxCanc-0.30031)*(SxCanc-0.30031)	-6.808917	2.358826	-2.89	0.0058*
Maint	0.5452336	0.194319	2.81	0.0072*
(%Studs-0.50031)*(SxCanc-0.30031)	-3.613176	1.313848	-2.75	0.0084*
(FltDur-1.75077)*(SxCanc-0.30031)	-2.407713	0.876783	-2.75	0.0085*
(FltDur-1.75077)*(Maint-1.00154)	-0.625548	0.282355	-2.22	0.0315*
(Lambda-0.06502)*(%Studs-0.50031)	-13.06033	22.26475	-0.59	0.5602
InBrf	0.0142581	0.196925	0.07	0.9426

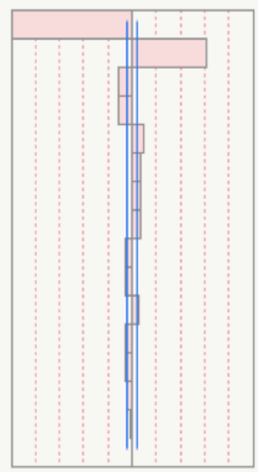


Figure 17. Summary of fit and sorted parameter estimates for OLS model with Max_A/CInUse as MOE.

This model satisfies the assumptions of an OLS model, as revealed in Figure 18. This figure shows the residuals by predicted plot, which implies constant variance. It also shows the qq norm plot, which suggests the errors are normal. A goodness-of-fit test concludes that H_0 cannot be rejected because it results in a p-value of 0.47.

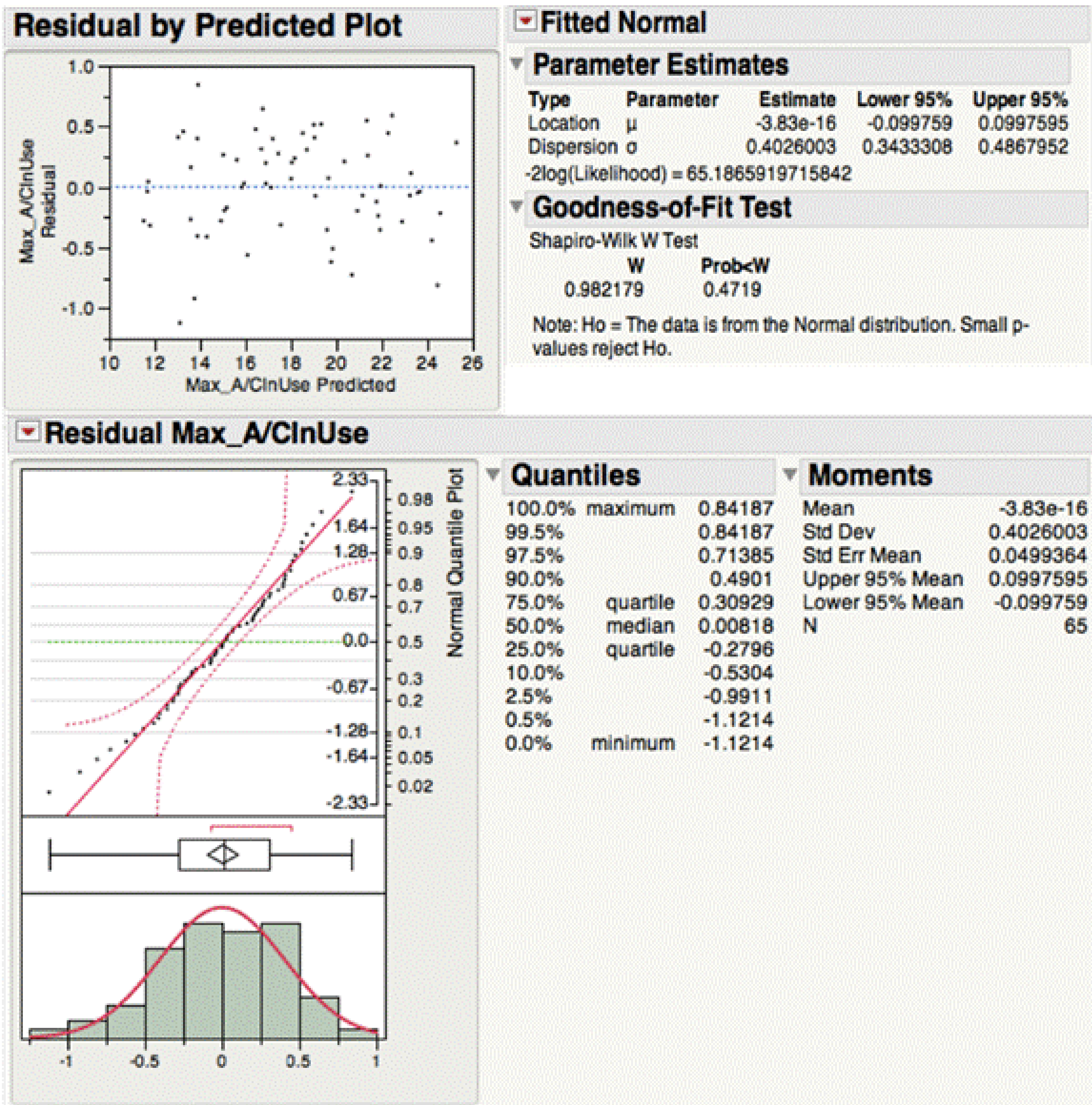


Figure 18. Residuals vs. predicted plot, qq norm plot of residuals with histogram, and goodness-of-fit test.

For 75.7% of the replications, Max A/C In Use is less than or equal to 20 A/C. This is important because normal usage rates will not allow more than 20 A/C to fly, and most likely something less than this. The reality is that it is impossible to use more A/C than the squadron is assigned, so Mean_DWQ discussed in this chapter are low estimates, and larger delays are expected.

F. CONCLUSIONS

The replica model demonstrates that weather is important in that it makes an impact on the Mean_DWQ. The replica model correctly replicates the ITC model's Flight Events. The two models observed react similarly under similar points of interest. The two models agree with importance of terms, and both models suggest that a significant amount of time is added to either TTT or DWQ as weather cancellation rates increase.

The thesis also examines resource utilization rates. These rates will become more significant as a comparison is done between the replica Simkit model and the improved Simkit model discussed in Chapter IV.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. IMPROVED SIMKIT MODEL

A. SIMKIT MODEL DESIGN

The improved Simkit model is similar to the replica model, but the one primary difference is the way weather is modeled. Instead of using independent Bernoulli trials to simulate the effect of weather, a stochastic process approach is used. A simple, continuous-time Markov process that alternates between good and bad weather is used. In essence, good weather is a resource that has limited availability. When it is available, all flights may use it; but when it is unavailable, all flights must wait until the weather changes.

The improved Simkit model flowchart, shown in Figure 19, illustrates the differences between the replica model and the improved model. The weather check occurs in the same place as the replica model. If it is bad weather, every student that comes into the queue goes directly into the weather wait queue, and when the weather changes from bad to good, all students are released from the wait queue. All students are checked to see if they have enough time to fly that day. If not, they go into an overnight wait queue and they have to recheck weather upon their release in the morning.

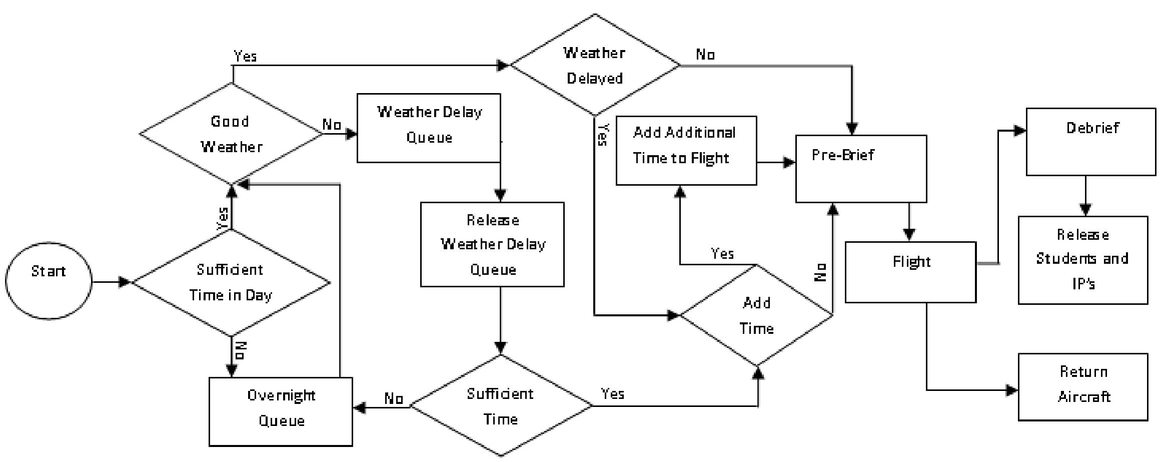


Figure 19. Flowchart of the improved Simkit model.

The improved Simkit model uses 11 different Java classes. Appendix C shows the event graph diagrams for this model. Most of them are the same as the Simkit replica model. This section explores the differences between them.

The first change is the Season Changer. In this model, SeasonChangerV2 puts the weather into four seasons, vice two. The duration of the seasons are varied in the DOE1.csv spreadsheet. Three of the four season durations are input as number of days in the design spreadsheet. The fourth is found by taking 360 days and subtracting the other three seasons' duration times.

Next is WheatherChanger, which uses a Markov process to transition from good to bad weather and back again. The duration of good and bad weather during two of the seasons are created using exponential distributions with Lambda rates, thus creating a continuous-time Markov process (Ross, 2007). During the other two seasons, the DOE1.csv spreadsheet gives parameters that increase or decrease these Lambda values.

The WaitCollectorV2 is similar to the WaitCollector; however, there is no longer an independent Bernoulli trial to check the weather. All flights coming into the weather checker during bad weather are delayed at least until the weather changes back to good. All students who go through the weather wait queue check to see if additional time is required, as explained in Chapter III.

Finally, the ITCModelExeV2 class works the same as the ITCModelExe. There are a few basic changes to the listener events, but essentially it controls the entire process.

B. SCOPE, LIMITATIONS, AND ASSUMPTIONS

The same limitations and assumptions carry into this model from the replica model. An additional assumption has to do with the weather. This thesis relies upon weather data that is not location-specific, as found in the DOE1.csv. It allows for the duration of the seasons and the duration of bad and good weather to fluctuate in the design. This construct is flexible enough to be able to model the weather in nearly any geographic location.

C. DESIGN OF EXPERIMENTS (DOE)

There are many similarities between the DOE found in Chapter III and the DOE here in Chapter IV. The MOE is the same. The list of variables is slightly different, as shown in Table 7. There is an increased number of Noise Variables and, instead of summer and winter cancelation rates, we now have good and bad weather durations.

Table 7. List of Decision and Noise Variables in the improved Simkit model.

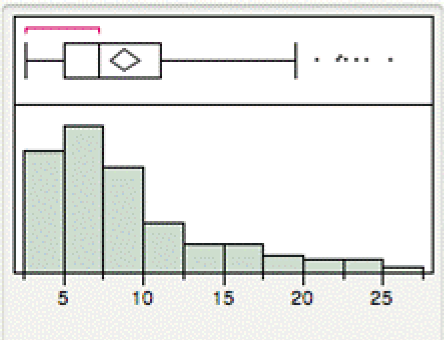
Decision Variables	**Noise Variables**
Arrival Rates (Lambda0)	In-Brief Times
Number of Students per Flight	Debrief Time
Flight Duration	Aircraft Flight Preparation Time
IP Break Time	Postflight Preparation Time
	Maintenance Time
	Additional Time
	Good Weather Duration (Lambda1)
	Bad Weather Duration (Lambda2)
	Winter Duration Adjustment (delta1)
	Winter Duration Adjustment (delta2)
	Summer Duration Adjustment (delta3)
	Summer Duration Adjustment (delta4)
	Spring Duration (time1)
	Summer Duration (time2)
	Fall Duration (time3)

The 19 factors listed in Table 7 are used in a $(N_0)_{22}^{129}$ NOLH design, which increases the number of design points to 129. Each design point is replicated 50 times, for a total of 6,450 runs. The seeds used with this model are the same as the ones used in the replica model. This use of common random number seeds reduces the variations caused by simulation (Law & Kelton, 2000). The entire experiment runs in approximately 25 minutes on a desktop computer with a Pentium 4 central processing unit, 3.6 gigahertz speed, and 3.0 gigabytes of random access memory. The results are collected using the same collection process found in Chapter III. The max pairwise correlation of coefficient estimate is 0.65.

D. ANALYSIS

The output of this model is the same as the output from Chapter III, except for the design points, which are slightly different. Figure 20 is a histogram of the primary response variable. Mean DWQ is 8.8 hours, with a fairly high Standard Deviation of 5.2 hours. All mean delay times fall at or below 25.5 hours. The fact that the mean weather delay went down is a bit surprising.

Mean_DWQ



Quantiles

100.0%	maximum	25.512
99.5%		25.512
97.5%		23.2197
90.0%		17.139
75.0%	quartile	11.0224
50.0%	median	7.1868
25.0%	quartile	5.0456
10.0%		3.476
2.5%		2.689
0.5%		2.6322
0.0%	minimum	2.6322

Moments

Mean	8.8265287
Std Dev	5.2689017
Std Err Mean	0.4639009
Upper 95% Mean	9.7444359
Lower 95% Mean	7.9086214
N	129

Figure 20. Distribution of the response variable Mean_DWQ.

The important factors in this model are found using a partition tree. There are 11 splits required before anything other than a property of weather enters the model. The first six splits are shown in Figure 21. An R^2 of about 0.8 is obtained with these six splits. Weather factors account for the majority of the variation in the response variable.

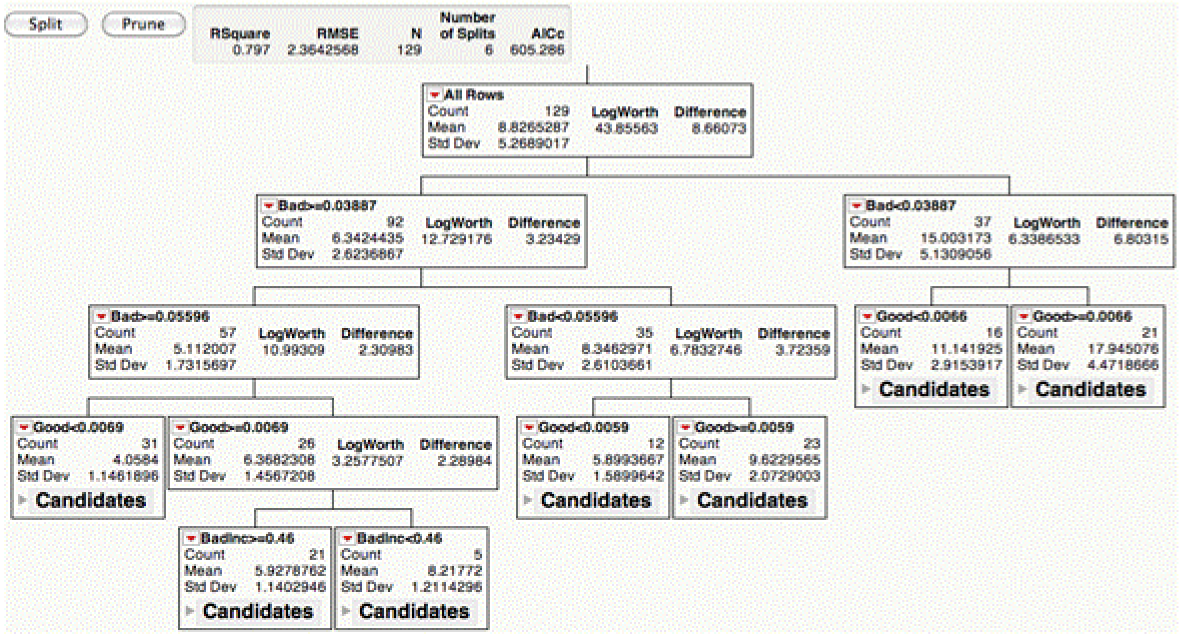


Figure 21. Partition tree with Mean_DWQ as a response variable.

An OLS model is fit to the new data using all 19 factors, with all two-way interactions. Stepwise AICc forward is used to help determine the factors that enter the model. This results in an adjusted R^2 of about 0.963, as shown in Figure 22. It is found that 16 of the original 19 factors, and 21 interaction terms, for a total of 37 terms out of the 191 possible, are statistically significant.

Summary of Fit	
RSquare	0.973578
RSquare Adj	0.963239
Root Mean Square Error	1.01021
Mean of Response	8.826529
Observations (or Sum Wgts)	129

Figure 22. Summary of fit for OLS model using Mean_DWQ as MOE.

To understand the importance of each factor, we examine the sorted parameter estimates shown in Figure 23. Good and Bad in Figure 23 refers to Lambda1 and Lambda2 from Table 7. Also, BadInc and GoodDec refer to delta2 and delta1 from Table7. It is clear that weather affects are extremely important in this model. The additional time shows that it increases the overall DWQ by an average of 7.5 hours.

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Bad	-226.7507	4.892512	-46.35	<.0001*
Good	1028.735	47.77992	21.53	<.0001*
BadInc	-10.53002	0.764538	-13.77	<.0001*
AddTime	7.4746922	0.76441	9.78	<.0001*
(GoodInc-0.45016)*(SpringDur-85.0078)	-0.948728	0.109949	-8.63	<.0001*
(SpringDur-85.0078)*(Summer Dur-80.031)	0.0043021	0.000502	8.57	<.0001*
(FltDur-1.75039)*(DeBrf-3.00155)	1.9567334	0.316947	6.17	<.0001*
(Bad-0.05205)*(BadInc-0.60031)	295.72659	49.7674	5.94	<.0001*
Summer Dur	0.044181	0.007644	5.78	<.0001*
GoodDec	1.860534	0.38232	4.87	<.0001*

Figure 23. Sorted parameter estimates of the OLS (10 of 37 variables are shown).

E. VALIDATION

1. Mean DWQ

This model satisfies the necessary OLS assumptions for a model. Figure 24 shows the residuals by predicted plot, which is not as clear as the previous plots. There seems to be a little shape to the points, but not enough to throw out the assumption of constant variance. The errors seem normally distributed, as shown by the qq-plot and the p-value from the goodness-of-fit test is around 0.94. Both of these are found in Figure 24 (Hamilton, 1992).

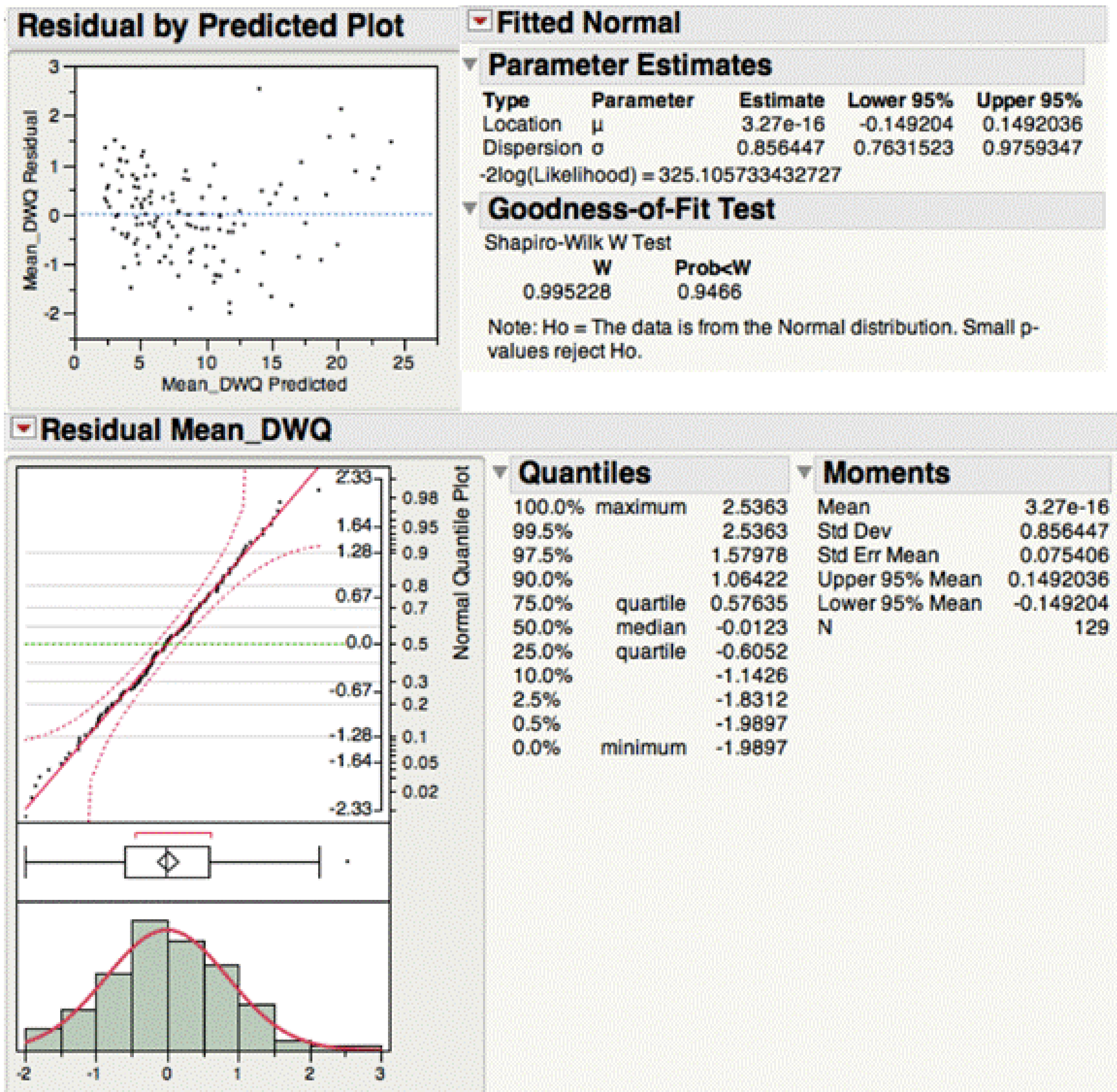


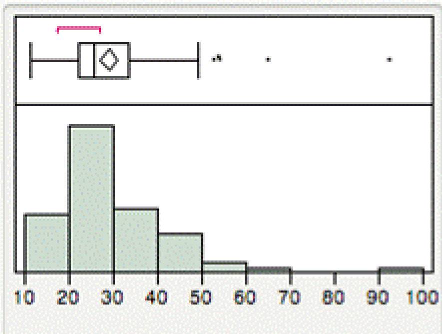
Figure 24. Mean_DWQ Residuals by Predicted Plot, QQ normal plot of residuals with histogram and goodness-of-fit test.

2. A/C Utilization

Figure 25 gives the distribution of the Mean (aircraftInUse_max), which is the mean of the maximum number of A/C flown at each repetition by the 129 design points. Appendix F shows the distributions of the Max A/C, Max brief rooms, and the Max IPs in use over the 6,450 data points. As is evident, the number of A/C utilized using this approach is significantly higher than the replica model. This suggests that wait times are

more affected by resources than with the replica model. Mean A/C utilization is around 30 A/C, vice a mean of 18 A/C from the previous model.

Max_A/CInUse



Quantiles

100.0%	maximum	92.56
99.5%		92.56
97.5%		54.1
90.0%		44.52
75.0%	quartile	33.42
50.0%	median	25.76
25.0%	quartile	21.96
10.0%		18.04
2.5%		12.87
0.5%		11.16
0.0%	minimum	11.16

Moments

Mean	28.922481
Std Dev	11.310088
Std Err Mean	0.9957977
Upper 95% Mean	30.892837
Lower 95% Mean	26.952125
N	129

Figure 25. Distribution of the Max_A/CInUse by design point.

Next, an OLS model is fit using the Max_A/CInUse, by design point, which results in an adjusted R^2 of 0.9533 (shown in Figure 26). The sorted parameter estimates are shown in Figure 27. The %Studs worked the same in this model as in the previous replica model.

Summary of Fit

RSquare	0.966474
RSquare Adj	0.953356
Root Mean Square Error	2.442673
Mean of Response	28.92248
Observations (or Sum Wgts)	129

Figure 26. Summary of fit for the Max_A/CInUse OLS model.

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
%Studs	-18.65051	0.739487	-25.22	<.0001*
Bad	-294.6647	11.83009	-24.91	<.0001*
Lambda	334.43108	14.78799	22.62	<.0001*
AddTime	-32.42984	1.84827	-17.55	<.0001*
BadInc	-20.54	1.848543	-11.11	<.0001*
(AddTime-0.50006)*(Bad-0.05205)	1087.1499	120.4692	9.02	<.0001*
(%Studs-0.50016)*(AddTime-0.50006)	47.507691	7.202556	6.60	<.0001*
(DeBrf-3.00155)*(Post-0.50008)	-13.41778	2.963797	-4.53	<.0001*
(Bad-0.05205)*(BadInc-0.60031)	518.71997	121.3392	4.27	<.0001*
(Lambda-0.06501)*(BadInc-0.60031)	-532.1887	125.2698	-4.25	<.0001*

Figure 27. Sorted parameter estimates Max_A/CInUse
(10 of the 37 variables shown).

This model also satisfies the OLS assumptions, as shown in Figure 28. The residuals by predicted plot implies constant variance. The qq norm plot had only minor granularity and some slight heavy tails, suggesting some outliers, but the errors seem normal (Hamilton, 1992).

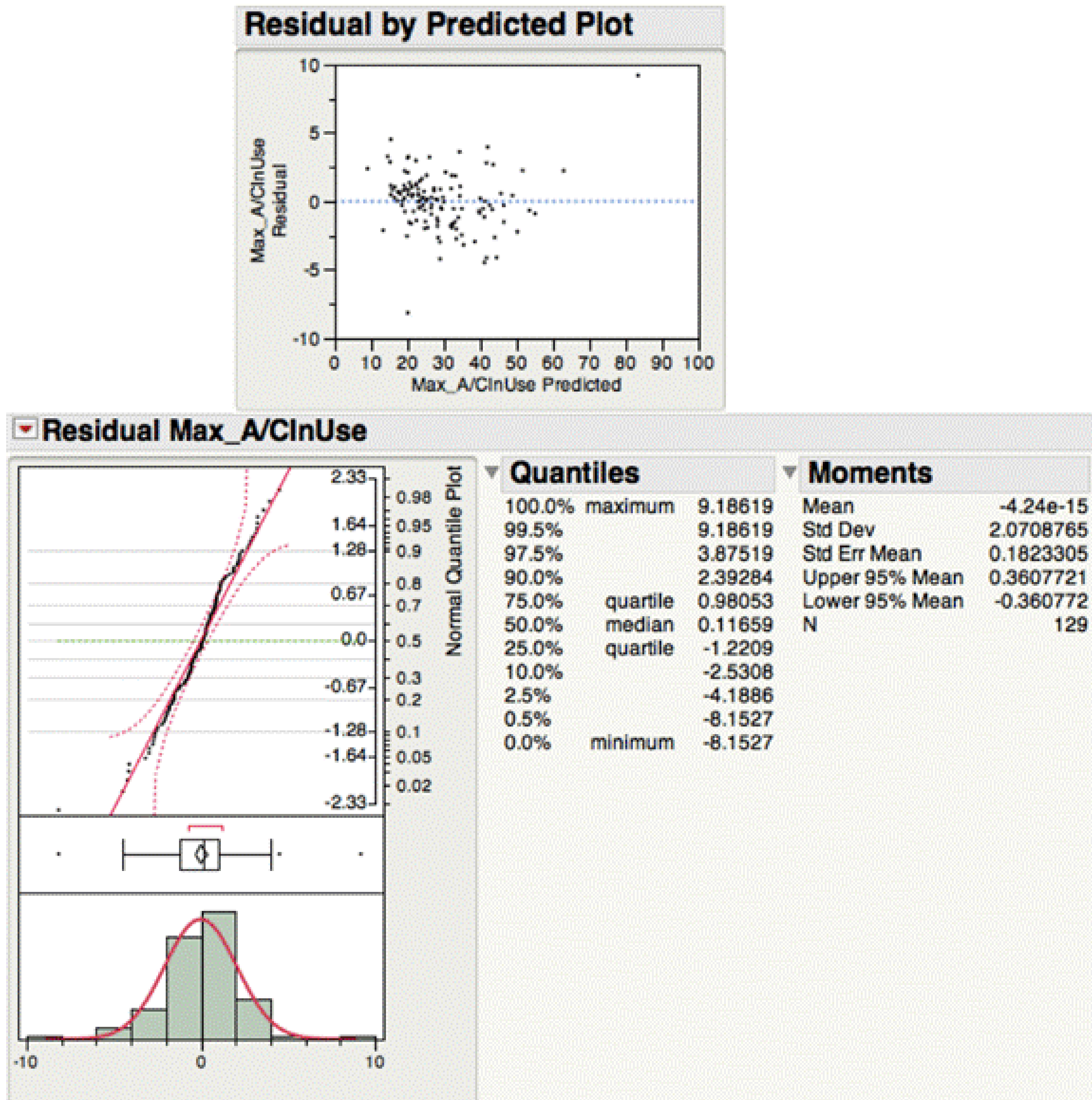


Figure 28. Residual by predicted, QQ norm plot, and goodness-of-fit test.

The extreme Max number of A/C in use is surprising, so a contour plot is created to get a better understanding of what the surface looks like. The contour plot

(Figure 29) shows that for a majority of the time, fewer than 30 A/C are used under normal operations. The normal expected %Studs, according to the current syllabi, is greater than 0.7. The contour plot clearly shows that when %Studs is above 0.7 the system typically uses less than 30 A/C and that the numbers fall closer to the 20 A/C assigned to the squadron.

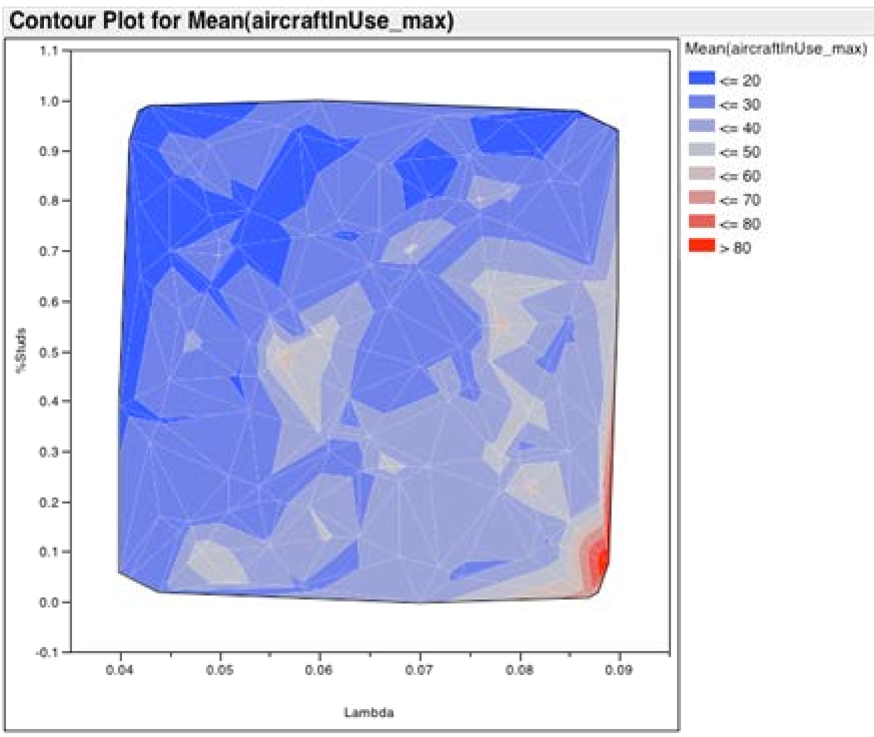


Figure 29. Contour plot for Mean of the Max_A/C against %Studs and arrival Lambda.

F. MODEL COMPARISON REPLICA SIMKIT VS. IMPROVED SIMKIT

1. DOE for the Comparison

One of the most obvious ways in which the ITC model diverges from reality is the assumption that weather affects all flights independently. Under this assumption, it is possible for two flights to arrive within minutes of each other, and for one to experience good weather and the other to experience a delay. To properly examine the effect of this assumption on the response variables of interest, it is necessary to set up a design that

allows the improved model to have the same weather cancellation rates as the replica model. Table 8 shows the design for the replica model.

Table 8. Replica model design.

WinterCanx	SummerCanx	All Other Factors' Default	Weather Canx Rate
0.1	0.1	See Table 10	0.1
0.2	0.2	See Table 10	0.2
0.3	0.3	See Table 10	0.3
0.4	0.4	See Table 10	0.4
0.5	0.5	See Table 10	0.5
0.6	0.6	See Table 10	0.6

To compare the two models against each other, the design uses a single rate for weather cancellations, as seen in Tables 8 and 9. In the long run, a weather cancellation rate for the replica model of, say, 0.2, will result in approximately 20% of all flights being delayed due to weather. For the improved model, by manipulating Good_Lambda and Bad_Lambda, we can achieve similar ratios of bad weather to good weather. For example, the second row of Table 9 results in bad weather approximately 20% of the time.

It is not clear, however, what the average duration of the good or bad weather should be. We code a value known as base, to facilitate comparison along this margin. For example, with a base of 50 hours, a ratio of 10 hours of expected bad weather and 40 hours of expected good weather yields a long-term weather cancellation rate of 0.2. Similarly, a ratio of 20 hours of expected bad weather and 80 hours of expected good weather also yields an approximate weather cancellation rate of 0.2. We select base levels of 50, 100, 150, and 200 hours. While time prevents us from determining the correct base, we are able to characterize its effect on the result.

Finally, all effects of seasonal changes were removed for the purposes of this experiment. The only difference between the simulations based upon this design is the way weather affects flights. The replica model uses a Bernoulli trial, while the improved model uses a continuous-time Markov process.

Table 9. Improved model design. (Base is the base number of hours for weather.)

Good Lambda	Bad Lambda	Base	All Other Factors' Default	Hours Bad	Weather Canx Rate	Hours Good
0.022222222	0.2	50	Tables 10 & 11	5	0.1	45
0.025	0.1	50	Tables 10 & 11	10	0.2	40
0.028571429	0.066666667	50	Tables 10 & 11	15	0.3	35
0.033333333	0.05	50	Tables 10 & 11	20	0.4	30
0.04	0.04	50	Tables 10 & 11	25	0.5	25
0.05	0.033333333	50	Tables 10 & 11	30	0.6	20
0.011111111	0.1	100	Tables 10 & 11	10	0.1	90
0.0125	0.05	100	Tables 10 & 11	20	0.2	80
0.014285714	0.033333333	100	Tables 10 & 11	30	0.3	70
0.016666667	0.025	100	Tables 10 & 11	40	0.4	60
0.02	0.02	100	Tables 10 & 11	50	0.5	50
0.025	0.016666667	100	Tables 10 & 11	60	0.6	40
0.007407407	0.066666667	150	Tables 10 & 11	15	0.1	135
0.008333333	0.033333333	150	Tables 10 & 11	30	0.2	120
0.00952381	0.022222222	150	Tables 10 & 11	45	0.3	105
0.011111111	0.016666667	150	Tables 10 & 11	60	0.4	90
0.013333333	0.013333333	150	Tables 10 & 11	75	0.5	75
0.016666667	0.011111111	150	Tables 10 & 11	90	0.6	60
0.005555556	0.05	200	Tables 10 & 11	20	0.1	180
0.00625	0.025	200	Tables 10 & 11	40	0.2	160
0.007142857	0.016666667	200	Tables 10 & 11	60	0.3	140
0.008333333	0.0125	200	Tables 10 & 11	80	0.4	120
0.01	0.01	200	Tables 10 & 11	100	0.5	100
0.0125	0.008333333	200	Tables 10 & 11	120	0.6	80

All the other factors are held constant at default values for both models. The default values are shown in Table 10. Lambda0 is the arrival rate, and %Stud is the number of students flying single or in a group of two. Flight Duration, Brief Time, DeBrief Time, IP Break, Pre Flight, Post Flight, and Maint Time are all in hours. ProbAdd is placed at 0.0, which essentially drops the additional time that is added in the previous runs.

Table 10. List of variable defaults that both models have in common.

Lambda0	%Stud	Flight Duration	Brief Time	DeBrief Time	IP Break	Pre Flight	Post Flight	Maint Time	Prob Add
0.065	0.7	1.8	1.75	3.0	1.0	0.5	0.5	1.0	0.0

In this thesis, it is desired to see what differences there are in the model strictly due to weather changes. The additional factors found in the improved model are set to ensure that they do not impact the improved model, as seen in Table 11. The variable names and functions are from Table 7. Deltas are values from 0.0–1.0 and show a percentage of change; 1.0 means no change. Times are in days.

Table 11. Variable defaults found only in the improved model.

Delta1	Delta2	Delta3	Delta4	Time1	Time2	Time3
1.0	1.0	1.0	1.0	90	90	90

For a proper analysis of variance it is necessary to use the same population in different experimental designs, which is what is done with this DOE (Devore, 2009). Weather is the only point of variation between the two models. The weather rates range from 0.1-0.6 in both models, while the response variables remain the same.

2. Analysis of Variance (ANOVA) Tests

Using the new design, the model's output is combined into one set of data using the concatenate function in JMP. An indicator variable, Model, denotes the model type. For example, a "1" refers to the replica model and a "2" refers to the improved model. Another indicator variable, Base, refers to the number of hours used as the base. A "1" in the base refers to the replica model, while all other bases refer to the improved model and its base number of hours.

A one-way ANOVA, with base as a factor and Mean DWQ as the response variable is shown in Figure 30. The replica model is assigned Base = 1; while the improved model has four levels as shown in Table 9. The highly significant F-statistic indicates differences between the models.

Summary of Fit

Rsquare	0.254264
Adj Rsquare	0.252269
Root Mean Square Error	16.50303
Mean of Response	18.2177
Observations (or Sum Wgts)	1500

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Base	4	138825.13	34706.3	127.4327	<.0001*
Error	1495	407163.15	272.3		
C. Total	1499	545988.28			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1	300	6.4374	0.95280	4.568	8.306
50	300	9.8280	0.95280	7.959	11.697
100	300	17.1236	0.95280	15.255	18.993
150	300	25.2230	0.95280	23.354	27.092
200	300	32.4766	0.95280	30.608	34.346

Std Error uses a pooled estimate of error variance

Figure 30. ANOVA of the two different models by Base Value.

Figure 30 only shows that there are differences, not which levels are different. A Tukey honestly significant difference (HSD) procedure (Devore, 2009) reveals significant differences in Mean DWQ between all but the replica model and the lowest base level of the improved model. The results are shown in Figure 31.

Comparisons for all pairs using Tukey-Kramer HSD

q*	Alpha
2.73115	0.05

Abs(Dif)-HSD

	200	150	100	50	1
200	-3.680	3.574	11.673	18.968	22.359
150	3.574	-3.680	4.419	11.715	15.105
100	11.673	4.419	-3.680	3.615	7.006
50	18.968	11.715	3.615	-3.680	-0.289
1	22.359	15.105	7.006	-0.289	-3.680

Positive values show pairs of means that are significantly different.

Level					Mean
200	A				32.476600
150		B			25.222967
100			C		17.123567
50				D	9.828000
1				D	6.437367

Levels not connected by same letter are significantly different.

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
200	1	26.03923	1.347467	22.3591	29.71936	<.0001*
200	50	22.64860	1.347467	18.9685	26.32873	<.0001*
150	1	18.78560	1.347467	15.1055	22.46573	<.0001*
150	50	15.39497	1.347467	11.7148	19.07510	<.0001*
200	100	15.35303	1.347467	11.6729	19.03316	<.0001*
100	1	10.68620	1.347467	7.0061	14.36633	<.0001*
150	100	8.09940	1.347467	4.4193	11.77953	<.0001*
100	50	7.29557	1.347467	3.6154	10.97570	<.0001*
200	150	7.25363	1.347467	3.5735	10.93376	<.0001*
50	1	3.39063	1.347467	-0.2895	7.07076	0.0875

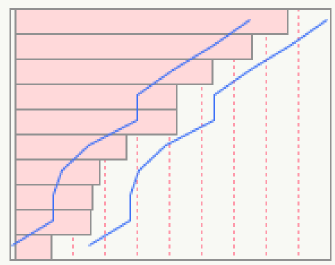


Figure 31. Part of the results found under the Tukey HSD procedure.

A boxplot graphically depicts the differences in the Mean DWQ in Figure 32. It clearly shows that as the base number of hours increases, the Mean DWQ increases as well. Using 150 hours as the base hours for weather will more than double the delay of the replica model. This large of a difference clearly demonstrates that the independence assumptions found in the ITC model are faulty and this results in artificially low TTT estimates.

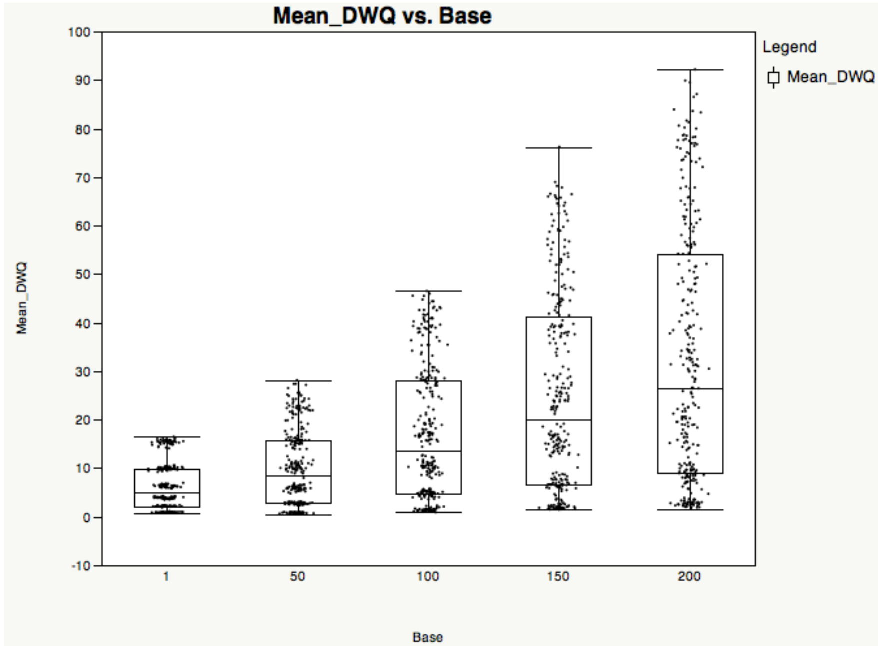


Figure 32. Boxplot of the Mean_DWQ at each Base level.

G. CONCLUSIONS

Three things are clarified in this chapter. First, weather is an important factor in the improved model. Second, aircraft utilization behavior indicates that the improved model likely underestimates DWQ due to the unconstrained resource assumption. Third, we find evidence that for the same long term probability of weather cancellations, the improved model estimates a higher DWQ than the replica model. Since the estimated DWQ for the improved model is likely a lower bound, it is plausible that the replica model (and by extension, the ITC model) substantially underestimates the contribution of weather delays in the total TTT. We find that the independent weather assumption in the ITC model is faulty and recommend modification.

IV. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

This thesis answers three questions. The first question is if weather is an important factor in the ITC model. Chapter II investigates the ITC model through DOEs and regression analysis. Summer and winter cancellation rates are some of the most important factors in the ITC model. This signifies that weather is important in the ITC model.

The next question is how sensitive TTT is with respect to weather. Chapter II also answers this question. We discover that a slight change in bad weather rates in the current model can result in substantial delays in TTT. For example, a small change in weather cancellation rates from 0.05–0.1 results in a full day of delay per pilot. Small changes in weather can drastically impact the TTT for each pilot.

The bulk of this thesis concentrates on the last question. If a change is made in the current model to more accurately reflect weather, are more accurate TTT estimates generated and are they statistically different than the current model? To answer this question, we create two new models—a replica model of the ITC model and an improved model, with proposed weather changes in it. These are necessary to quantify the differences that exist between the current way weather is modeled and the proposed approach.

In Chapter III, we analyze the replica model, which replicates the flight event from the ITC model. This is the only event that is affected by weather in the ITC model. To ensure that the ITC model and the replica model behave similarly, we compare the results of the two models. The same factors are important in the ITC model and the replica model. The two models also produce similar results in the number of pilots that train during similar durations of time. Resource limitations are not used in the replica model, which is one major difference between the two models. Analysis of the A/C usage allows decision makers to determine if the usage is acceptable or not.

The improved model found in Chapter IV is similar to the replica model, with an improved weather check. The replica model is modified in Simkit using a continuous-time Markov process to change states from bad to good and back again. The season changer uses four distinct weather durations instead of two.

The importance of weather in the improved model is evaluated using Mean DWQ and resource utilization rates. Mean DWQ is affected significantly by changes in weather patterns. Resource utilization rates increase significantly using the improved model. We demonstrate that these models are different both visually and graphically, using resource utilization rates and boxplots.

Quantifying the differences in these models is important so that decision makers know if these changes are worth spending money on. To quantify the results, a controlled DOE is run. We populate both models with identical input data and seed information. Weather is the only point of variation in the two models, and doing this reduces variances caused by simulation. The analysis evaluates the output of the two models using ANOVA techniques. We also use Student's t tests to establish that the models are different. The results of these tests validate that there are statistically significant differences in the two models. It proves that the independence assumption made in the ITC model is faulty. This implies that TTT estimates in the original ITC model are underestimated and improvements need to be made.

B. RECOMMENDATIONS

We show that the independent weather assumption in the ITC model is inappropriate. We recommend modifying the ITC model to reflect good weather as a resource, and that it is necessary to conduct a flight, which is only intermittently available. One way to do this is to use a continuous-time Markov process, as we demonstrate in this thesis.

C. SUGGESTIONS FOR FURTHER RESEARCH

One area for further research, directly related to the weather, is to fit the model to METAR data for Eglin AFB, Florida. Azimetli's thesis does this using METAR data for the air bases in Turkey.

Another area of analysis is to evaluate the effect that maintenance has on response variables in the current implementation. Also, it would be important to evaluate the effect of extending the ITC model to more realistically account for maintenance.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. ITC MODEL GROUND RULES AND ASSUMPTIONS

1. Normal operating hours for the Integrated Training Center are 0700-2300 Monday through Friday. Briefing rooms and equipment will be available to support a 0700 take off and commencement of simulator events at 0700. Simulation of a single shift operation can be accomplished by giving all students the same work day start time (i.e. 0800) with a work day length of 9 hours (workday ends at 1700).

2. Students are inducted evenly throughout the year and commence training on Monday of their induction week. The induction files can be adjusted to simulate a non-uniform flow of students into the ITC.

3. More than one class may be inducted in the same week.

4. The optimum pilot class size for JSF is 6 students. The current ITC classroom design for pilots can accommodate up to 12 students per class.

5. Maintainer class size is 12 students. The current ITC classroom design for maintainers can accommodate up to 16 students per class.

6. Pilot and Maintainer classes are service unique.

7. The first flight brief for the day occurs at 0445 for a 0700 takeoff.

8. The first flight simulator brief of the day occurs at 0600 for a 0700 simulator event.

9. The student workday is a 9.0 hour day (start to finish) and students are assigned a specific "start workday time". The intention of the 9.0 hour day is to give each student 8.0 hours per day when training can be accomplished and 1.0 hour per day for a meal break. Maintenance training for the U.S. Services is accomplished on two shifts.

10. In no case will students and instructors be scheduled for a training event sooner than 12.0 hours following the completion of their last event from the previous day. As an example, a student and/or instructor who finishes a flight debrief at 2200 will not be scheduled for a training event before 1000 the next day.

11. Maximum crew duty day for instructor pilots is 12.0 hours measured from the start of the first training event to engine shutdown. This ground rule is based on a United States Air Force Instruction 11-202 Volume 3 General Flight Rules, "For single seat aircraft or when only one pilot has access to the flight controls, the maximum flight duty period is 12 hours." The normal instructor work day is 9.0 hours.

12. Aircraft System Maintenance Trainer events contain 12 students and 2 instructors.

13. Ejection System Maintenance Trainer events contain 4 students and 1 instructor.

14. Weapons Load Trainer events contain 4 students and 1 instructor.

15. Engine module replacement training utilizes training devices provided by Pratt & Whitney & General Electric with a ratio of 1 instructor for every 4 students.

16. A Full Mission Simulator event and flight event will be permitted on the same day for the same student if they can both be accomplished within the students 9.0 hour work day.

17. Flight simulator set up time is 15 minutes for each training event (after initial daily startup).

18. Students will be permitted to fly more than one flight event on the same day. This only occurs for short flights such as Field Carrier Landing Practice events due to work day length restrictions.

19. Student planning time for each flight event is 1.0 hour utilizing the Mission Planning Station.

20. Brief time for training flights is 2.25 hours prior to takeoff.

21. Flight debriefs end 3.0 hours after landing.

22. Day flights land before 1900 and night flights takeoff after 1900.

23. Airborne aggressor requirements are satisfied by Training Squadron JSF aircraft flown by Training Squadron instructor pilots.

24. After landing, there is a 10% probability the aircraft will require repair. Time to repair is 3.0 hours followed by a 1.5 hour preparation time for the next flight event. These parameters are adjustable via the user interface.

25. Aircraft turnaround time for the next flight event is 1.5 hours and is adjustable via the user interface.

26. Cancellations for weather occur after student planning and prior to commencement of the flight brief. Weather cancellation rates are 5% and are adjustable via the user interface.

27. Re-fly rates for flights and simulator events are 8% and 5% respectively and are adjustable via the user interface. The re-fly decision is made at the end of debrief.

28. Navy Carrier Qualification (CQ) re-fly rate is 12.5%. Students who repeat CQ will re-fly all simulator and flight events associated with the CQ phase of training.

29. Students will not re-fly the same flight or simulator event more than once.

30. Student planning time for each Full Mission Simulator event is 30 minutes and is adjustable via the user interface. Student planning utilizes a Mission Planning Station.

31. Full Mission Simulator brief time is 60 minutes and is adjustable via the user interface.

32. Full Mission Simulator debrief time is 30 minutes and is adjustable via the user interface.

33. The number of Military instructor pilots is equal to the number of training aircraft in accordance with direction received from the Services and the JSF Program Office. All airborne flight instruction is provided by Military instructor pilots.

34. When determining maintainer instructor quantities, actual on-task training time for each instructor is limited to approximately 30 hours per week. This allows 10 hours per week of instructor time for student counseling, syllabus review, course content review, lesson preparation, and completion of other instructor duties.

35. After completing a flight debrief, the instructor will not be scheduled for a subsequent flight brief for 1.0 hour.

36. On each training day, there is a 2% probability that each instructor and student may be unavailable for a 2.0 hour time period. This is intended to account for medical appointments and unexpected absences for both instructors and students.

37. Minimum allowable spacing between flights on the same low-level route is 10 minutes.

38. Landing aircraft have priority over departing aircraft for the runway.

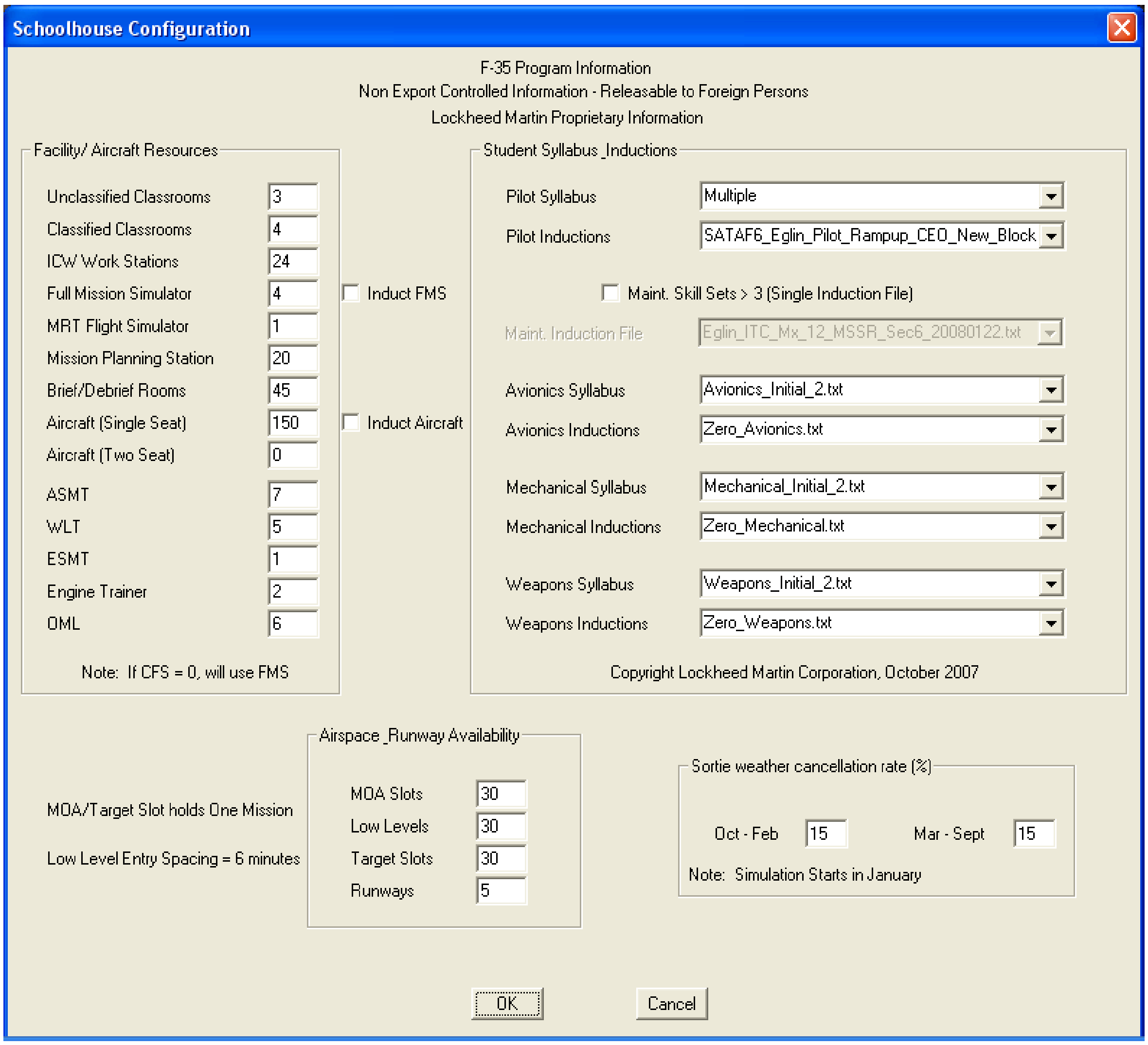
39. There are a total of 246 days per year that are available for. This parameter is selectable via the user interface.

40. Students will take a 15-minute break in between non-like training events (e.g. simulator to classroom, classroom to simulator, etc.). Breaks during successive classroom training events will be handled by the instructor (Kenney, 2009, pp. 8-10).

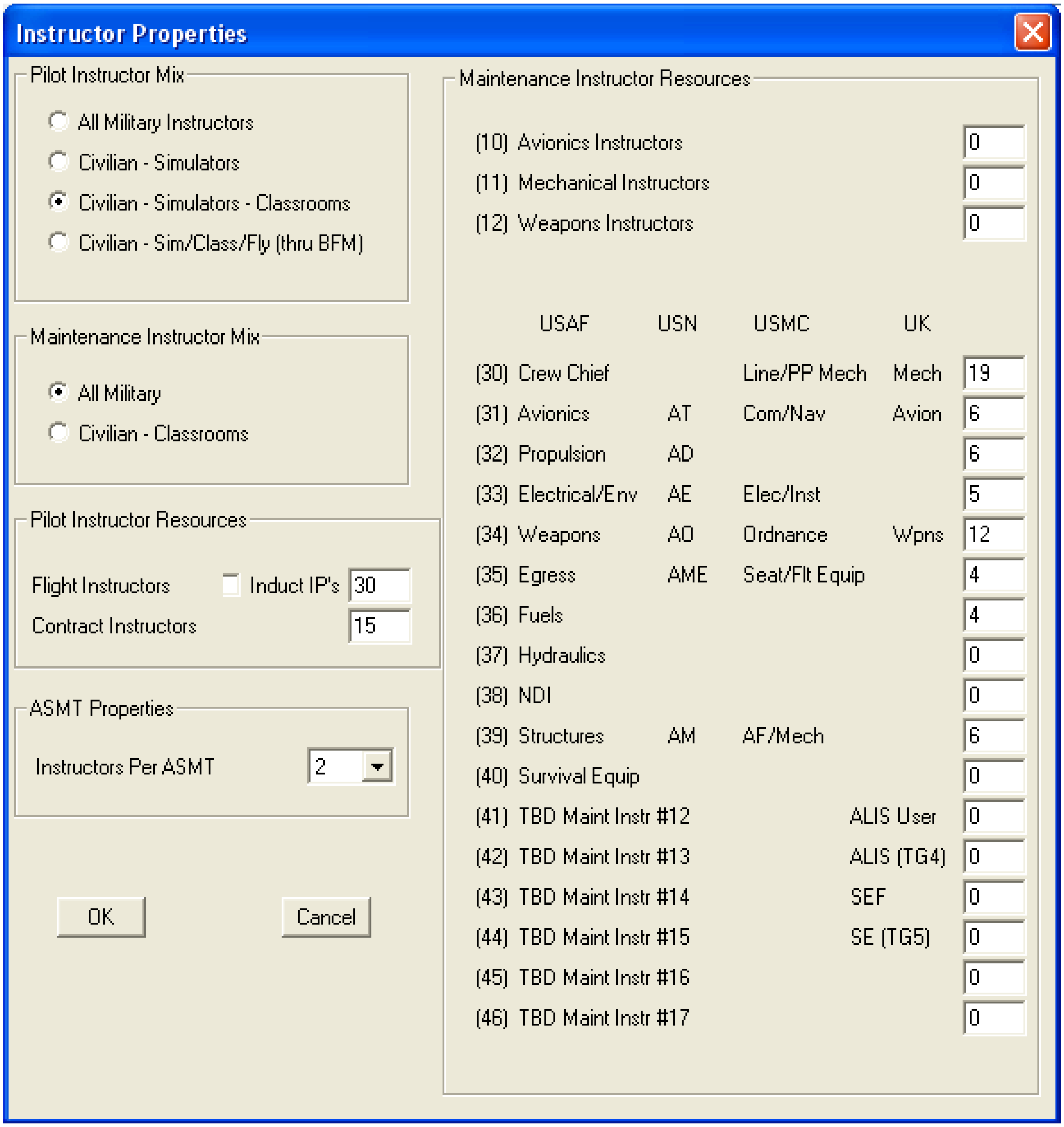
THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. ITC MODEL GUI SCREENSHOTS

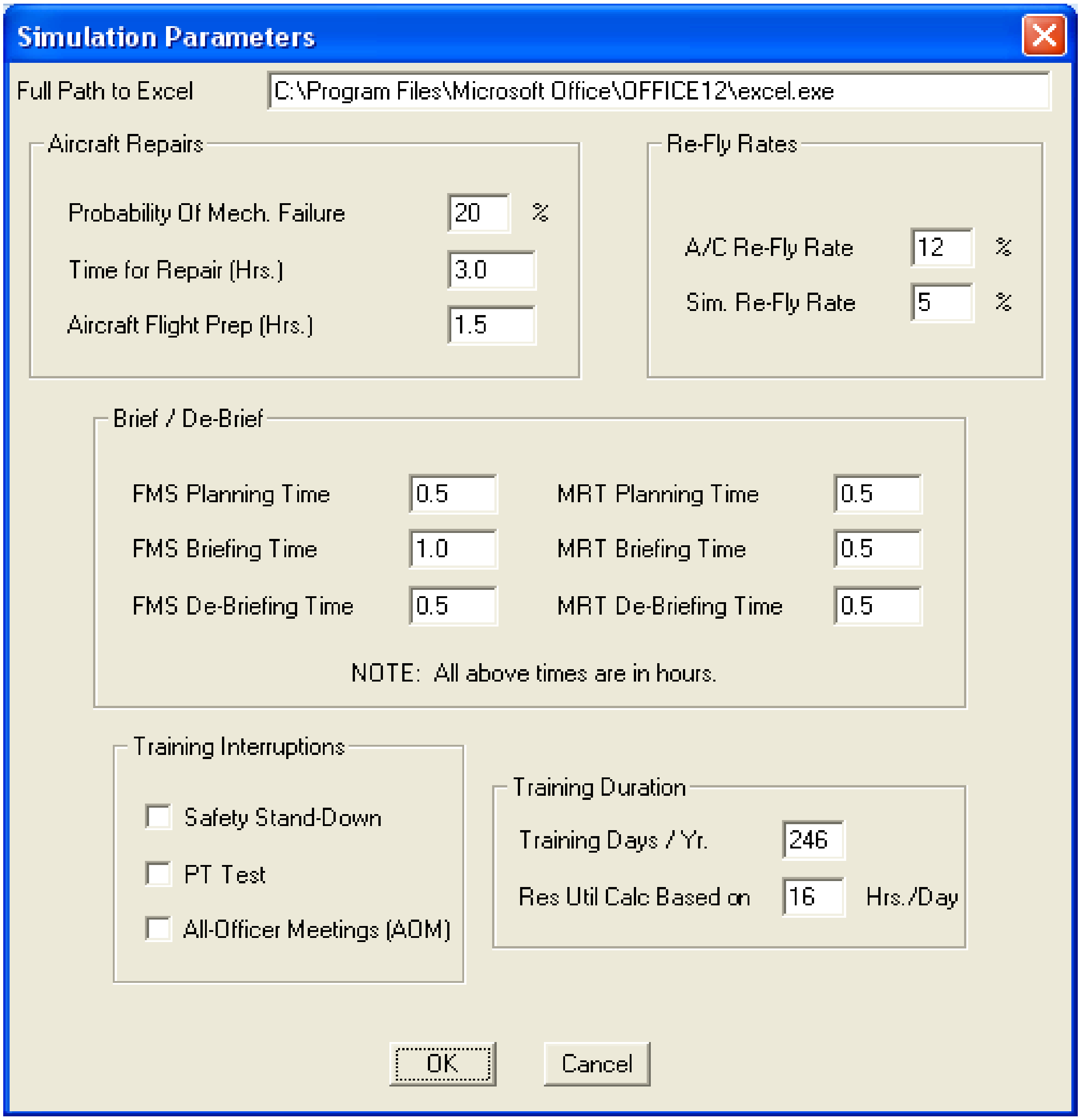
1. Below is the Scenario User Interface, where the user inputs the pilot induction schedules and various factors such as Classrooms available (Unclassified, Classified, Brief Rooms, FMS, etc.), Runways available, and Weather Cancellation percentages.



2. This is the Instructor User Interface and allows the user to input the number of IPs available, Number of Civilian Instructors available, and how the instructors will be used. It also allows the user to input the Maintenance resources, which are only important if the maintenance piece is on. It will not be used in this thesis.

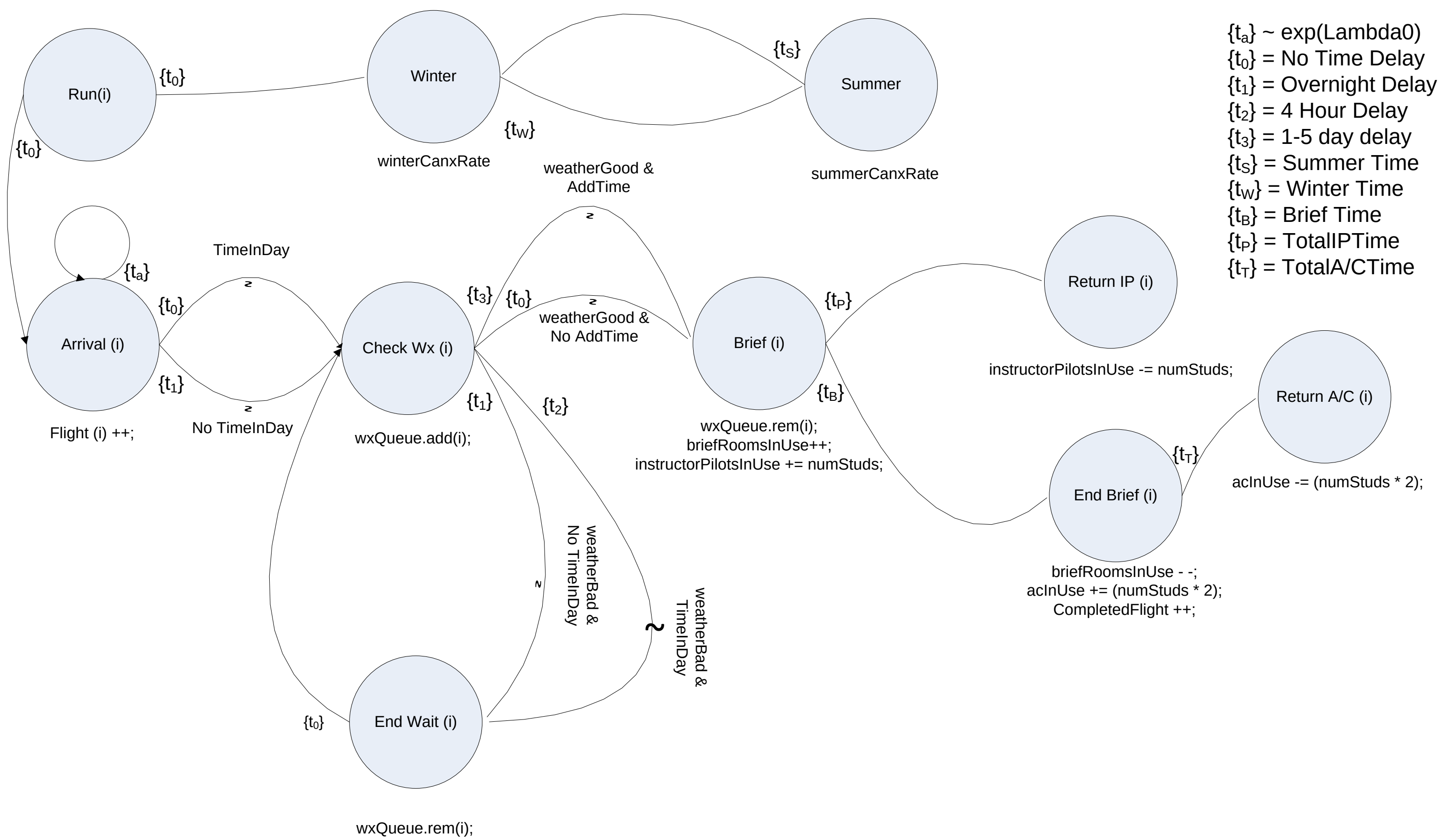


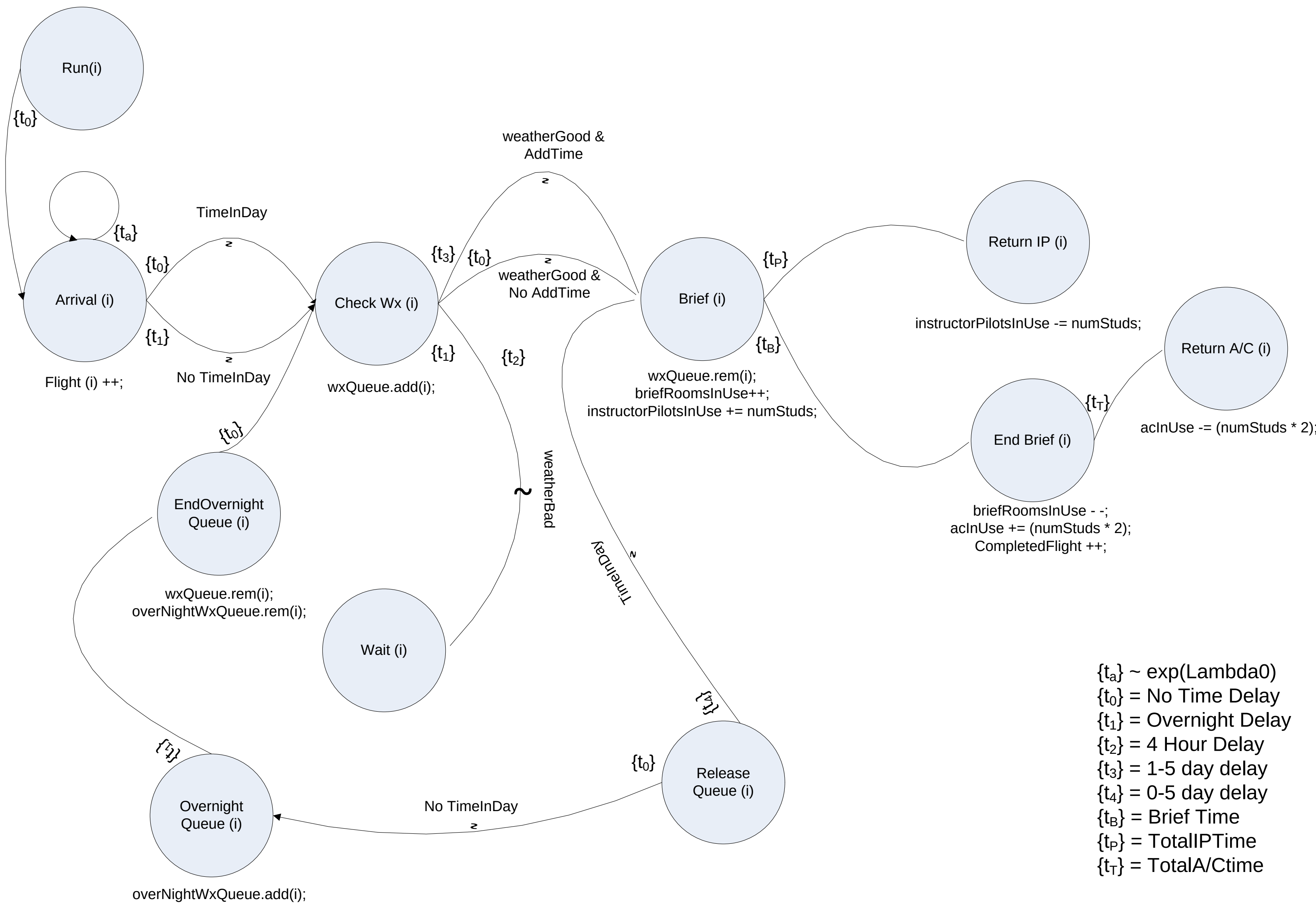
3. Finally, the Attribute User Interface. This is where the user is allowed to manipulate factors such as Probability of Mechanical Failure, repair times, number of A/C and Sim refly rates, length of the different briefs involved pre and post flight, the number of training days available to fly in, and the number of hours available each day to train.



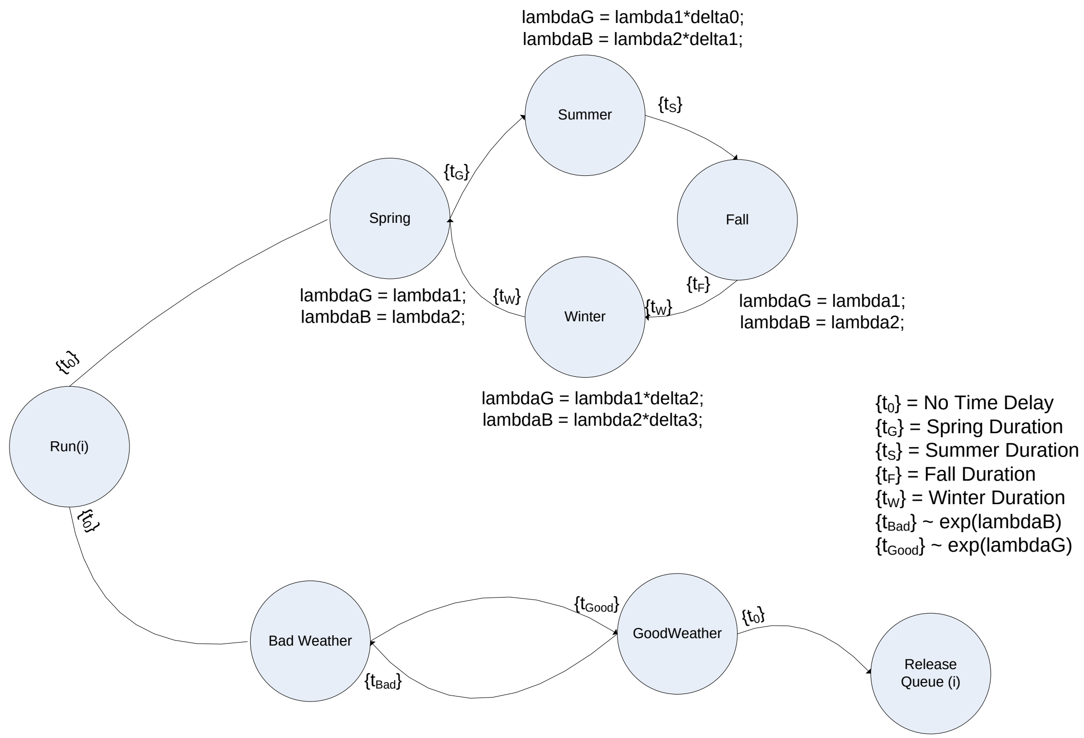
THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C. EVENT GRAPH DIAGRAMS FOR THE REPLICA MODEL AND THE IMPROVED MODEL



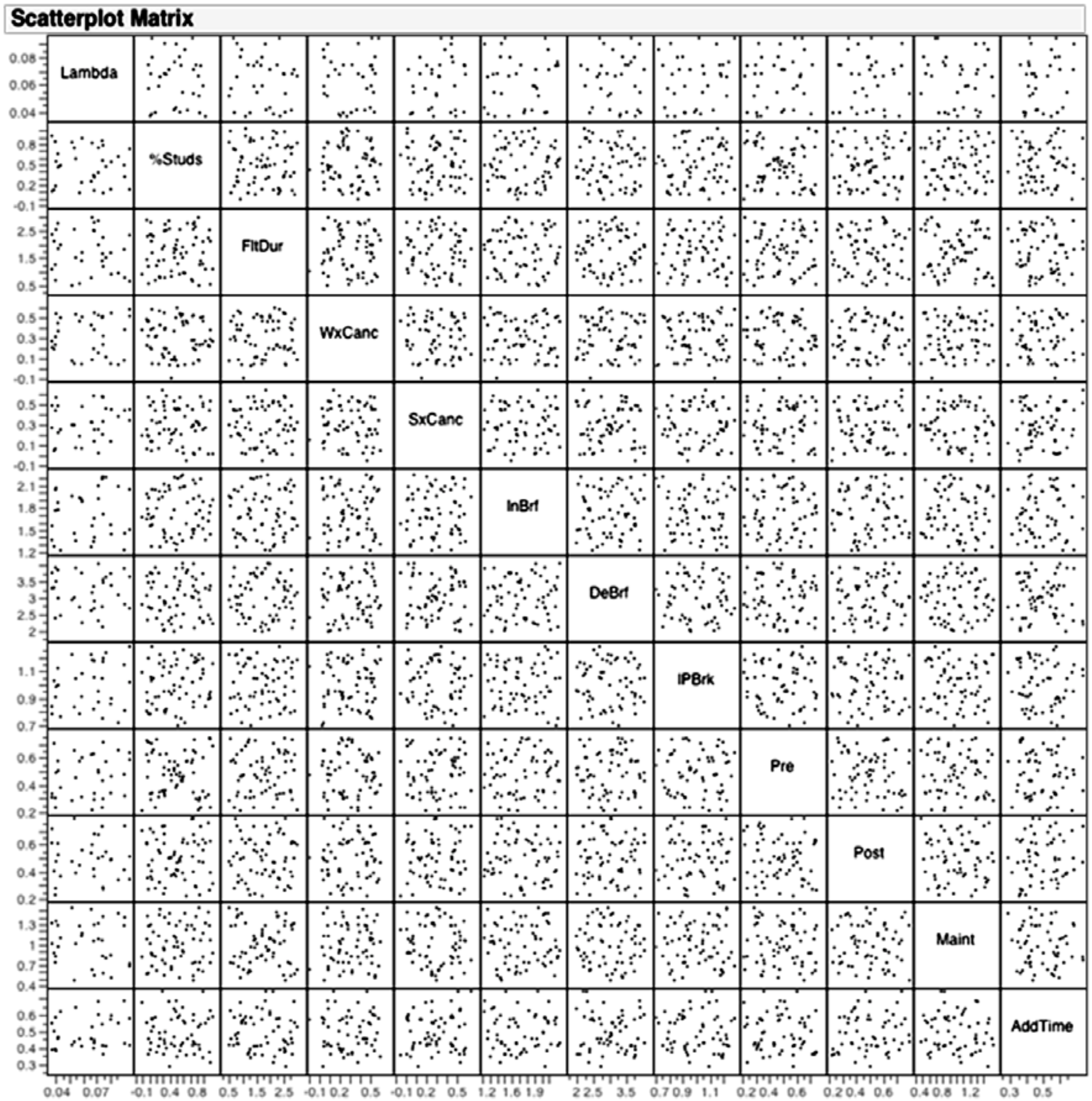


{ta} ~ exp(Lambda0)
{t0} = No Time Delay
{t1} = Overnight Delay
{t2} = 4 Hour Delay
{t3} = 1-5 day delay
{t4} = 0-5 day delay
{tB} = Brief Time
{tP} = TotalIPTime
{tT} = TotalA/Ctime


lambdaG = lambda1*delta0;
lambdaB = lambda2*delta1;
Summer
{t_S}
{t_G}
Spring
Fall
lambdaG = lambda1;
lambdaB = lambda2;
{t_W}
Winter
{t_F}
{t_W}
lambdaG = lambda1;
lambdaB = lambda2;
{t_0}
lambdaG = lambda1*delta2;
lambdaB = lambda2*delta3;
Run(i)
{t_0}
{t_{Good}}
GoodWeather
{t_0}
Bad Weather
{t_{Bad}}
Release Queue (i)
{t_0} = No Time Delay
{t_G} = Spring Duration
{t_S} = Summer Duration
{t_F} = Fall Duration
{t_W} = Winter Duration
{t_{Bad}} ~ exp(lambdaB)
{t_{Good}} ~ exp(lambdaG)

THIS PAGE INTENTIONALLY LEFT BLANK

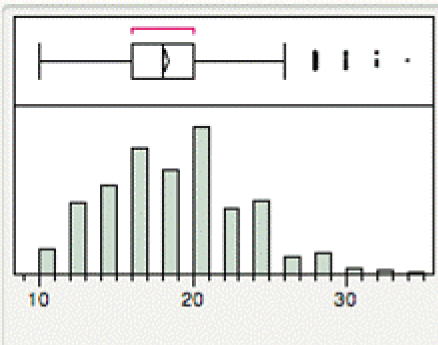
APPENDIX D. CHAPTER III SPACE-FILLING PROPERTIES



THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX E. CHAPTER III RESOURCE UTILIZATION BAR GRAPHS WITH ASSOCIATED QUANTILES AND MOMENTS

aircraftInUse_max



Quantiles

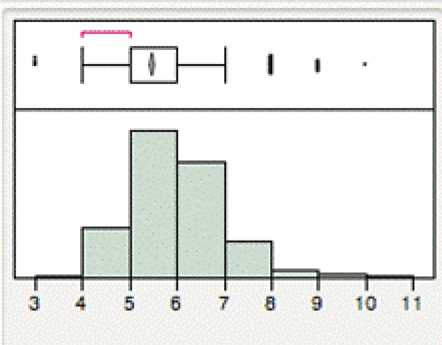
100.0%	maximum	34
99.5%		30
97.5%		28
90.0%		24
75.0%	quartile	20
50.0%	median	18
25.0%	quartile	16
10.0%		12
2.5%		10
0.5%		10
0.0%	minimum	10

Moments

Mean	18.217231
Std Dev	4.3755186
Std Err Mean	0.0767517
Upper 95% Mean	18.367717
Lower 95% Mean	18.066744
N	3250

Distribution of the A/C utilization rates from the 3,250 design points.

briefRoomsInUse_max



Quantiles

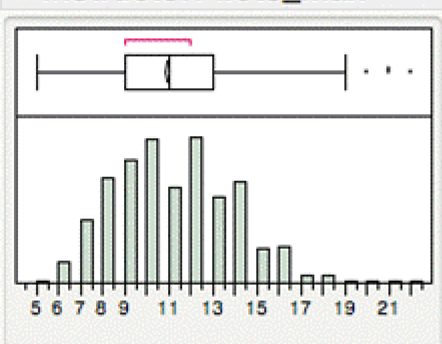
100.0%	maximum	10
99.5%		8
97.5%		7
90.0%		7
75.0%	quartile	6
50.0%	median	5
25.0%	quartile	5
10.0%		4
2.5%		4
0.5%		4
0.0%	minimum	3

Moments

Mean	5.4566154
Std Dev	0.9431979
Std Err Mean	0.0165448
Upper 95% Mean	5.4890547
Lower 95% Mean	5.4241761
N	3250

Distribution of the brief room utilization rates from the 3,250 design points.

instructorPilots_max



Quantiles

100.0%	maximum	22
99.5%		18
97.5%		16
90.0%		14
75.0%	quartile	13
50.0%	median	11
25.0%	quartile	9
10.0%		8
2.5%		7
0.5%		6
0.0%	minimum	5

Moments

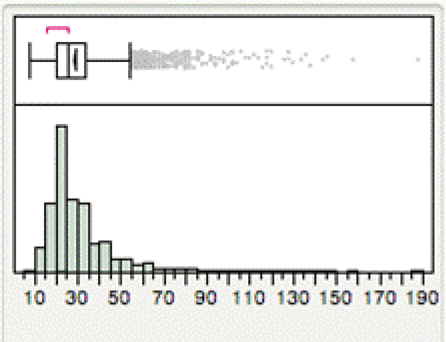
Mean	10.972615
Std Dev	2.6077614
Std Err Mean	0.0457432
Upper 95% Mean	11.062304
Lower 95% Mean	10.882927
N	3250

Distribution of the IPs' utilization rates from the 3250 design points.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX F. CHAPTER IV RESOURCE UTILIZATION BAR GRAPHS WITH ASSOCIATED QUANTILES AND MOMENTS

Max_A/CInUse



Quantiles

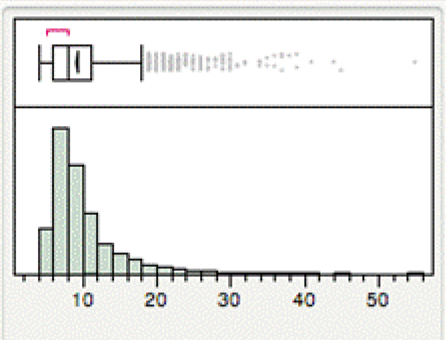
100.0%	maximum	188
99.5%		100
97.5%		66
90.0%		46
75.0%	quartile	34
50.0%	median	26
25.0%	quartile	20
10.0%		16
2.5%		14
0.5%		10
0.0%	minimum	8

Moments

Mean	28.922481
Std Dev	14.210626
Std Err Mean	0.176943
Upper 95% Mean	29.269348
Lower 95% Mean	28.575614
N	6450

Distribution of the A/C utilization rates from the 6,450 design points.

Max_BriefRoomsInUse



Quantiles

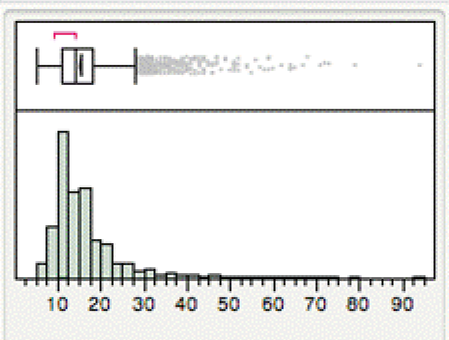
100.0%	maximum	55
99.5%		29
97.5%		21
90.0%		14
75.0%	quartile	11
50.0%	median	8
25.0%	quartile	6
10.0%		5
2.5%		5
0.5%		4
0.0%	minimum	4

Moments

Mean	9.191938
Std Dev	4.2889578
Std Err Mean	0.0534038
Upper 95% Mean	9.2966271
Lower 95% Mean	9.0872489
N	6450

Distribution of the brief room utilization rates from the 6,450 design points.

Max_IPsInUse



Quantiles

100.0%	maximum	94
99.5%		50.745
97.5%		34
90.0%		23
75.0%	quartile	18
50.0%	median	14
25.0%	quartile	11
10.0%		9
2.5%		7
0.5%		6
0.0%	minimum	5

Moments

Mean	15.310543
Std Dev	7.0783991
Std Err Mean	0.0881364
Upper 95% Mean	15.483319
Lower 95% Mean	15.137766
N	6450

Distribution of the IPs' utilization rates from the 6,450 design points.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Axtell, R., Axelrod, R., Epstein, J. M., & Cohen, M. D. (1995, September 1). Aligning simulation models: A case study and results. *Computational and Mathematical Organization Theory*, *1*, 123–141.

Azimetli, M. (2008). Simulation of Flight Operations and Pilot Duties in LANTIRN Squadrons Using Simkit. Monterey CA: Naval Postgraduate School.

Buss, A. (2001). Simulation news Europe technical notes: Discrete event programming with Simkit. Monterey, CA: Naval Postgraduate School.

Cioppa, T. M. (2002). Efficient nearly orthogonal and space-filling experimental designs for high-dimensional complex models. Monterey, CA: Naval Postgraduate School.

Cioppa, T. M., & Lucas, T. W. (2007). Efficient nearly orthogonal and space-filling Latin hypercubes. *Technometrics*, *49*(1), 45.

Devore, J. L. (2009). *Probability and statistics for engineering and the sciences* (7th ed.). C. Crockett (Ed.). Belmont, CA: BROOKS/COLE CENGAGE Learning.

MOE, (1998). DoD modeling and simulation (M&S) glossary. Washington, D.C.: Office of the Under Secretary of Defense (Acquisition and Technology). This document cleared for public release (Distribution A) by DoD Office of Security Review (Case No. 10-S-2163) March 2010.

Goble, J. (1997). MODSIM III–A tutorial. *Proceedings of the 1997 Winter Simulation Conference*. S. Andradóttir, K. J. Healy, D. H. Withers, & B. L. Nelson (Eds.). La Jolla, CA: CACI Products Company.

Hamilton, L. C. (1992). *Regression with graphics: A second course in applied statistics*. M. J. Sugarman (Ed.). Belmont, CA: Duxberry Press.

Kenney, P. (2009). Joint strike fighter integrated training center simulation model accreditation support package (Integrated training center development for SDD). Tampa: Paul Kenney.

Kenny, P. (2010, May 15). OAD JSF. P. Kenny (Ed.). Retrieved from https://cle.nps.edu/xsl-portal/site/abaa373f-4c32-4dde-a6e0-4a8e740bef7f/page/6b82168a-80fb-492f-90ea-ba110c73a453

Kovach, G. C. (2010, December 8). Hard choices ahead for Marines. signonsandiego.com. Retrieved from http://www.signonsandiego.com/news/2010/dec/08/commandant-calls-joint-strike-fighter-essential/

Law, A. M., & Kelton, W. D. (2000). *Simulation modeling and analysis* (3rd ed.). E. M. Kevin, & T. Kane (Eds.). San Francisco, CA: Thommas Casson.

Lucas, T. (2010). Improving analysis with the *integrated* training center model for time-to-train estimates. Naval Postgraduate School, Monterey, CA: SEED Center.

Ross, S. M. (2007). *Introduction to probability models* (9th ed.). T. Singer (Ed.). Berkley, CA: Academic Press.

Rabachault, M. (2011, March 1). F-35 looking more like white elephant. Yahoo! News. Retrieved from http://news.yahoo.com/s/afp/20110113/pl_afp/usmilitaryaerospacef35_20110113153609

Sanchez, S. M. (2008). Better than a petaflop: The power of efficient experimental design. *Proceedings of the 2008 Winter Simulation Conference* (S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, & J. W. Fowler (Eds.). Monterey, CA: Naval Postgraduate School.

Sanchez, S. M. (2005, May 3). NOLHDesigns_v4, xls: Generating nearly orthogonal Latin hypercube designs. Monterey, CA: Naval Postgraduate School.

Sanchez, S. M. (2000). Robust design: Seeking the best of all possible worlds. *Proceedings of the 2000 Winter Simulation Conference* (J. A. Joines, R. R. Barton, K. Kang, & P. A. Fishwick (Eds.) Monterey, CA: Naval Postgraduate School.

SAS Institute Inc. (2009). JMP 8 statistics and graphics guide (2nd ed.). Cary, NC: SAS Institute Inc.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Marine Corps Representative
   Naval Postgraduate School
   Monterey, California

4. Director, Training and Education
   MCCDC, Code C46
   Quantico, Virginia

5. Director, Marine Corps Research Center
   MCCDC, Code C40RC
   Quantico, Virginia

6. Marine Corps Tactical Systems Support Activity
   (Attn: Operations Officer)
   Camp Pendleton, California

7. Director, Studies and Analysis Division
   MCCDC, Code C45
   Quantico, Virginia

8. LtCol Luis Villalobo
   F-35 Lightning II Program Office
   Arlington, Virginia

9. Director, SEED Center
   Naval Postgraduate School
   Monterey, California

10. Maj Lieberman
    USMC AETC VMFAT- 501/S-3
    Eglin AFB, Florida

11. Capt Dummar
    Naval Postgraduate School
    Monterey, California